Check for updates

Adaptive Behavior

Original Paper

Adaptive Behavior
19(4) 250–263
© The Author(s) 2011
Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/1059712311411112
adb.sagepub.com
SAGE

Human sensorimotor learning for humanoid robot skill synthesis

Jan Babič[1], Joshua G. Hale[2] and Erhan Oztop[3]

Abstract

Humans are very skilled at learning new control tasks, and in particular, the use of novel tools. In this article we propose a paradigm that utilizes this sensorimotor learning capacity to obtain robot behaviors, which would otherwise require manual programming by experts. The concept is to consider the target robot platform as a tool to be controlled intuitively by a human. The human is therefore provided with an interface designed to make the control of the robot intuitive, and learns to perform a given task using the robot. This is akin to the stage where a beginner learns to drive a car. After human learning, the skilled control of the robot is used to build an autonomous controller so that the robot can perform the task without human guidance. We demonstrate the feasibility of this proposal for humanoid robot skill synthesis by showing how a statically stable reaching skill can be obtained by means of this framework. In addition, we analyze the feedback interface component of this paradigm by examining a dynamics task, in which a human learns to use the motion of the body to control the posture of an inverted pendulum that approximates a humanoid robot, so that it stays upright.

Keywords

Sensorimotor learning, humanoid robot, balanced reaching, posture control

1 Introduction

If robots were able to imitate human motion demonstrated to them, the acquisition of complex robot skills would be straightforward. Motion capture systems may be used to record the motion of a human subject and map it to the kinematic structure of a robot (Atkeson et al., 2000). Differences in the kinematic and dynamic properties of a humanoid robot and a human demonstrator complicate this process, and the efficacy of the approach depends on the mapping implemented by the researcher (Schaal, 1999). This approach can be considered an open-loop approach to robot skill generation, as illustrated in Figure 1. Most of the studies identifying with the headings ''robotic imitation'' or ''teaching by demonstration'' follow the open-loop approach (Bentivegna, Atkeson, & Cheng, 2004; Billard & Hayes, 1999; Billard & Siegwart, 2004; Breazeal & Scassellati, 2002; Ito & Tani, 2004; Schaal, 1999). In the classical method of direct teaching for robot motion synthesis (Kushida, Nakamura, Goto, & Kyura, 2001; Ude, Gams, Asfour, & Morimoto, 2010), the user moves the robot by actively holding and manipulating the robot through desired postures. Although this circumvents the problems of ad hoc mapping, it is hard to apply this technique to dynamic tasks such as full body humanoid walking and posture control. This approach may also be considered open-loop, because the human is not in the control loop of the robot. In this article, we propose a method which is similar to direct teaching, but also suitable for dynamic tasks. The method involves placing the human into the control loop, to act as an adaptive controller in order achieve a given task with the robot. This may be compared with teleoperation, in which the goal is to facilitate skillful control of a machine. The goal of the proposed method, however, is an autonomous controller derived from the human's performance where

[1]Jožef Stefan Institute, Ljubljana, Slovenia
[2]Cyberdyne Inc., Ibaraki, Japan
[3]NICT Advanced ICT Research Institute, Brain ICT Laboratory, ATR Cognitive Mechanisms Laboratories, and Osaka University, Japan

Corresponding author:
Jan Babič, Jožef Stefan Institute, Jamova cesta 39, SI-1000 Ljubljana, Slovenia
Email: jan.babic@ijs.si

remote manipulation and remote perception are not the critical issues (Chen, Haas, & Barnes, 2007). We refer this closed-loop approach as the *human learning for robot skill synthesis (HL-RSS) paradigm*. The key components of this paradigm are *human sensorimotor learning* and *autonomous robot operation*. Human learning here refers to the classical motor learning or sensorimotor learning (Wolpert & Ghahramani, 2000; Wolpert, Ghahramani, & Flanagan, 2001). Although this field has been well studied in behavioral and brain sciences using simple motor tasks such as reaching and pointing (Shadmehr & Wise, 2005), it has not been studied in the context of human control of robots, which needs to be addressed in the coming decades. The latter component utilizes comparatively well-established techniques from machine learning. Of the many inquiries concerning the HL-RSS paradigm, two key questions are (a) How well does it work, and can a human really control a humanoid robot in the same manner as he can control some common tools and machines in everyday life? and (b) Would a feedback interface that affects vestibular and proprioceptive senses enable the human to learn quicker and perform better than with the use of the visual feedback? In this article we address these questions with two separate robotic experiments. In the first experiment the subjects were free of learning dynamics relations (i.e., the task is velocity independent), whereas in the second experiment the subjects had to actively learn dynamics relations.

Our previous work with an anthropomorphic robotic hand indicated that the HL-RSS paradigm is a tenable method for obtaining manipulation tasks (Oztop, Lin, Kawato, & Cheng, 2006, 2007). In this study, we show that the paradigm is also applicable to humanoid robots by obtaining a balanced reaching skill, lending support to an affirmative answer to question (a). In the human learning phase of both the manipulation and the balanced reaching skills, the human subject was provided feedback in the visual domain. In the manipulation task of Oztop et al. (2006), the human subject directly observed the robot under control; whereas in the balanced reaching task, the subject was presented with an abstract but intuitive visualization expressing the balance of the robot (this study). In both cases, through visual feedback, the human subject was able to learn to perform the task with the robot, thus facilitating the extraction of autonomous controllers for these tasks. The ability of the human brain to control novel tools (Goldenberg & Hagmann, 1998; Imamizu, 2010) suggests that we can learn to control complex machines (cars, planes, and humanoid robots). Yet, the time it takes a human to master the control system may be critical. Thus far, we have seen that visual feedback is sufficient for humans to control robots skillfully when dynamics is negligible. Can humans learn tasks with nonnegligible dynamics, and if yes, what is the best feedback modality? To address these issues, we performed a separate second experiment to compare visual and proprioceptive-vestibular feedback, in which subjects controlled an inverted pendulum approximating a humanoid robot.

The article is organized as follows. Section 2 introduces the HL-RSS paradigm in a general setting, and

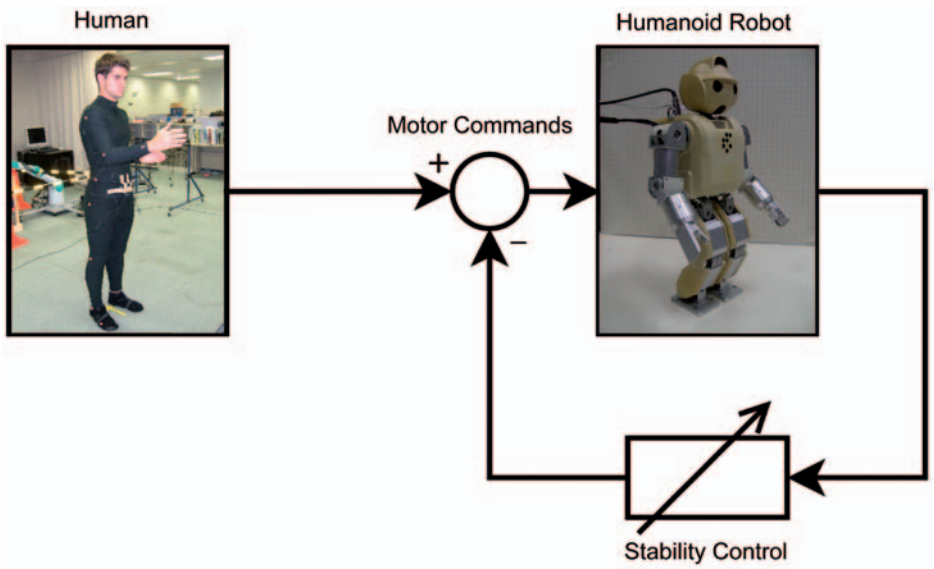

Figure 1. Open-loop control of the humanoid robot. When the motion of the human is transferred to the robot directly there is no guarantee that the robot will be stable unless a local controller takes care of the robot's postural stability.

Sections 3 and 4 explain how the balanced reaching skill is acquired according to the paradigm. After presenting the results of this endeavor (Section 5), the second experiment assessing the comparative effect of visual versus proprioceptive-vestibular feedback in controlling an inverted pendulum is described (Section 6) and the results presented (Section 7). Section 8 indicates future directions and discusses several issues related to the experiments conducted in this study. Finally, the experiments, results, and contributions of this study are summarized in Section 9.

2 Human Learning for Robot Skill Synthesis (HL-RSS) paradigm

The key idea of the HL-RSS paradigm is to exploit humans' natural proficiency for learning to use novel tools in order to obtain an autonomous motor controller for a robot which would otherwise require expert programming. The construction of the motor controller has two phases. In the first phase, a human controls the robot via an intuitive closed loop interface as illustrated in Figure 2 and trains to become proficient in the target skill with the robot. Although defining a metric for quantifying the load of the human in a human–robot interaction scenario is hard (Steinfeld et al., 2006), we anticipate that long-term training will not be uncommon for complex robot tasks. This is analogous to a beginner learning to drive a car; practice is essential. In the second phase, the skilled performance of the human is used to collect data points that may be exploited by a machine learning algorithm that acquires a *policy* for controlling the robot autonomously. This machine learning step effectively reverse engineers the control policy discovered by the human brain. This has parallels with the paradigm of direct teaching; but the number of data points extracted from the human performance is typically huge in comparison with direct teaching. From a biological perspective, in direct teaching, human cognition is at work in producing the data points, whereas in HL-RSS the human central nervous system (CNS), in particular the motor system, is at work in data generation. The CNS has evolved for motor control and is extremely adaptive. It incorporates dedicated circuits for the control and learning of the dynamic characteristics of the limbs and external objects (Imamizu, 2010; Shidara, Kawano, Gomi, & Kawato, 1993). The cognitive system on the other hand, is not oriented towards the minute-to-minute error correction and real-time feedback processing critical for control tasks.

3 Balanced reaching skill through HL-RSS

An attempt was made to equip a free-standing humanoid robot with the skill to reach targets using its hand without falling over. A small, 50 cm high humanoid robot from Fujitsu, HOAP-II, was adopted as the target platform. The robot has 25 rotational joints: two arms with 5 degrees of freedom, two legs with 6 degrees of freedom, the head with 2 degrees of freedom,

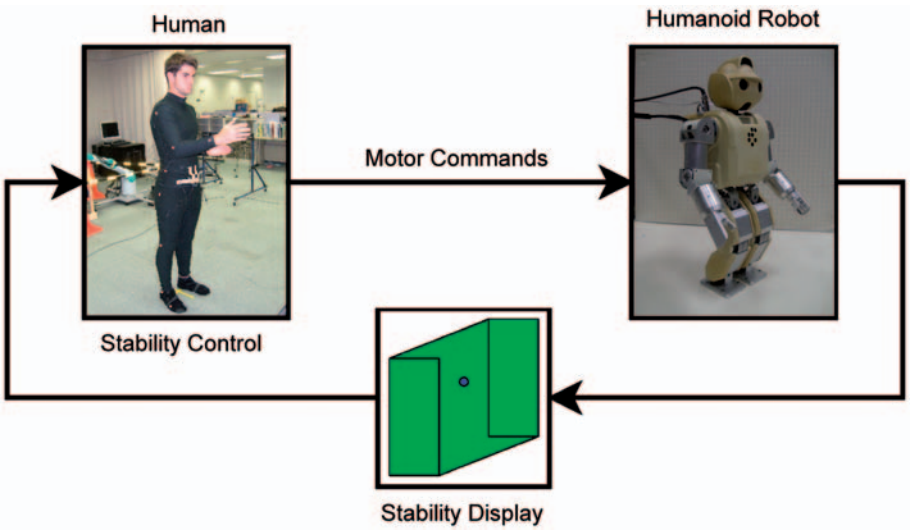

Figure 2. Closed-loop control of the humanoid robot. The motion of the human is transferred to the robot while the robot's stability is maintained *by the human* in the loop. The human learns to provide motor signals that satisfy both the task requirements and the constraints of remaining stable through a display which expresses the stability with a rendering of the robot's support polygon and its projected center of mass.

and the body with 1 degree of freedom. From a classical robotics perspective, the task is to solve the inverse kinematics problem with the postural stability constraint—an objective that has been widely studied. Here, however, our aim was to show that this skill can be obtained through the HL-RSS paradigm. We began by setting up the infrastructure for controlling the robot in real-time with human motion capture. As the task of the human would be to keep the robot stable while using the hand to reach a target specified by the experimenter, we also designed a stability feedback mechanism. In this setup, the motion of the human demonstrator was acquired using an active marker-based motion capture system (Visualeyez, PhoeniX Technologies Inc.). Sixteen markers were placed symmetrically on the demonstrator's body; specifically, over the little-toe joints, ankle, knee, hip, over the iliac crest, shoulder, ankle, and wrist joints. The joint configuration of the demonstrator's body was determined in real time using an exact analytical inverse kinematics formulation. Before the experiment, the subject stood still in the zero posture with arms straight down by the sides of the body and looking forwards. The joint configuration in this posture was compared with the joint configuration of the same posture of the humanoid robot. The difference between each joint angle was later used as the offset of the demonstrator's joint configuration. Joint range limits about 5° short of the physical limits of the humanoid robot were applied before the angles were sent to command the robot. There was no scaling of the joint angles. For the stability feedback, the position of the robot's center of mass projected onto the floor was rendered with a frequency of 50 Hz along with the robot's polygon of support. This visualization was presented to the demonstrator who was also prevented from seeing the robot during the experiment. The humanoid robot's center of mass was first analytically determined in its reference coordinate frame based on the current state of its body. To faithfully project the center of mass onto the floor even if the robot's feet were not flat on the floor, an inclinometer (XSens MTi, XSens), mounted at the back of the robot was used.

The demonstrator's task was to slowly sweep the sagittal plane with the right hand as directed by the experimenter and to keep the robot's center of mass within the support polygon. After a short practice session of about 15 minutes the demonstrator was able to move the body, albeit slowly, in such a way that the robot's center of mass remained within the support polygon. The robot was therefore statically stable during the motions generated by the demonstrator. The motions were imitated by the robot in real time, but could also be played back later, without the demonstrator. The placement of the robot's end effector was constrained to the sagittal plane (i.e., only upward/downward and forward/backward motion was permitted). The curve in Figure 3 shows the position of the robot's end effector during its imitation of the human. Each data point on the trajectory represents a reaching target position for the humanoid robot's end effector. The axes of the graph define the sagittal plane of the humanoid robot where the x-axis represents the horizontal axis and the y-axis represents the vertical axis of the plane. An average velocity of the humanoid robot's end effector during the motion generated by the demonstrator was 1.7 cm/s. A consequence of this rather slow motion of the robot's end effector is the jaggedness of the trajectory that can be observed in Figure 3.

The reaching motion of the humanoid robot obtained during its imitation of human motion shown in Figure 3 was *statically stable* (i.e., the projection of the robot's center of mass on the ground remained inside the robot's support polygon). To quantify the stability we defined a static balance metric as

$$stability = 1 - \frac{|x_{COM}|}{l_{foot}/2}, \tag{1}$$

where x_{COM} is the distance in the sagittal plane of the robot's center-of-mass projection from the center of the support polygon and l_{foot} is the length of the robot's

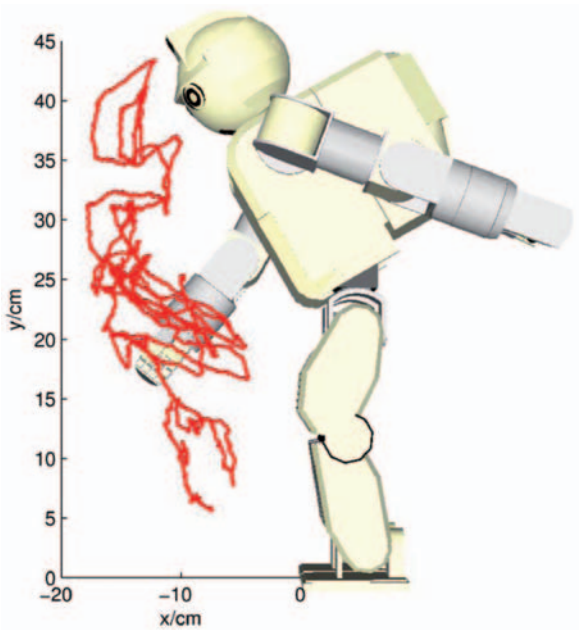

Figure 3. The trajectory of the humanoid robot's end effector obtained during motion transfer together with a ghost image of the reaching humanoid robot. Each data point of the trajectory represents the reaching target position of the humanoid robot's end effector.

foot. A plot of the static stability measure of the humanoid robot during its imitation of human motion is shown in Figure 4.

4 Statically stable inverse kinematics determination

While the robot end effector traced the trajectory shown in Figure 3 through human guidance, the robot joint angles and the end-effector positions were recorded. The key observation that leads to a statically stable inverse kinematics mapping is that the collected data came from a joint angle manifold that was guaranteed to provide static stability for the robot, as the robot was stable during the full data collection period. Therefore, joint angles picked from a faithful representation of this manifold will ensure that the robot will be stable with these angles. Statically stable inverse kinematics then can be achieved by finding data points on this manifold (i.e., joint angles) that will minimize the distance between the actual and the desired end-effector positions. There are indeed many methods that can be used for this purpose. In particular any powerful nonlinear regression method would do, such as Gaussian process regression, support vector machine (SVM) regression, and locally weighted projection regression (LWPR), to name a few. Here we choose a straightforward method, namely, radial basis function networks (RBFs), underlining the fact that in this study the emphasis is on the *data and its generation* rather than the method to represent it.

The relation between the joint angles and end-effector position is nonlinear due to the trigonometric functions involved in the forward kinematics mapping of a rotary robotic chain. RBFs can capture this nonlinearity as they are universal approximators, that is, they can represent any well-defined function to a desired accuracy with a finite number of basis functions (Park & Sandberg, 1991). As such, RBFs (as well as other powerful learning methods) are able to represent the joint constraints embedded in the data by the human demonstrator's performance.

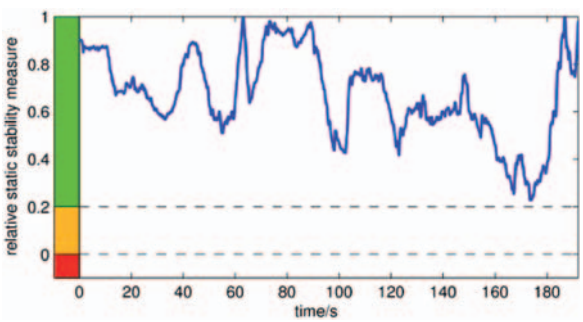

Figure 4. The stability of the humanoid robot during its imitation of human motion shown in Figure 3.

Assuming a total of m data points, we use m-row matrices $\mathbf{X}$ and $\mathbf{Q}$ to denote the humanoid robot end-effector positions and the corresponding robot joint angles respectively. We can consider the basis functions as the vector function Γ operating on the end-effector space defined as, $\Gamma(\boldsymbol{x}) = (\varphi_1(\boldsymbol{x}), \varphi_2(\boldsymbol{x}), \ldots, \varphi_N(\boldsymbol{x}))$ with

$$\varphi_i(\boldsymbol{x}) = e^{-\frac{|\boldsymbol{x}-\boldsymbol{\mu}_i|}{\sigma^2}}, \tag{2}$$

where $\boldsymbol{\mu}_i$ and σ^2 are open parameters to be determined. Effectively, $\Gamma(\mathbf{X})$ converts each row of $\mathbf{X}$ into an N-dimensional vector forming a new high-dimensional data matrix $\mathbf{Z}$:

$$\mathbf{Z} = \Gamma(\mathbf{X}) = \begin{bmatrix} \varphi_1(\boldsymbol{x}_1) & \varphi_2(\boldsymbol{x}_1) & \ldots & \varphi_N(\boldsymbol{x}_1) \\ \varphi_1(\boldsymbol{x}_2) & \varphi_2(\boldsymbol{x}_2) & \ldots & \varphi_N(\boldsymbol{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_1(\boldsymbol{x}_m) & \varphi_2(\boldsymbol{x}_m) & \ldots & \varphi_N(\boldsymbol{x}_m) \end{bmatrix}. \tag{3}$$

With N large enough (and with nondegenerate choices of $\boldsymbol{\mu}_i$ and σ^2), the relationship between $\mathbf{Z}$ and $\mathbf{Q}$ can be considered linear, that is, $\mathbf{Q} = \mathbf{ZW}$, where $\mathbf{W}$ is a real matrix that can be solved in the least squares sense with

$$\mathbf{W} = \mathbf{Z}^{+}\,\mathbf{Q}, \tag{4}$$

where $\mathbf{Z}^{+}$ represents the pseudo-inverse of $\mathbf{Z}$. The residual error is given by

$$tr\big((\mathbf{ZW} - \mathbf{Q})\,(\mathbf{ZW} - \mathbf{Q})^T\big). \tag{5}$$

Although residual error gives a measure of the linearity of the mapping, it cannot be used to determine N that will ensure good general robot performance. This requires cross-validation as will be detailed later on. Once $\mathbf{W}$ is computed, then for a given desired end-effector position $\boldsymbol{x}_{des}$, the joint angles $\boldsymbol{q}_{pred}$ that would bring the end effector to $\boldsymbol{x}_{des}$ are given by

$$\boldsymbol{q}_{pred}^T = \Gamma(\boldsymbol{x}_{des}^T)\,\mathbf{W} = \begin{pmatrix} \varphi_1(\boldsymbol{x}_{des}) & \varphi_2(\boldsymbol{x}_{des}) & \ldots \varphi_N(\boldsymbol{x}_{des}) \end{pmatrix}\mathbf{W}. \tag{6}$$

By using Equation 6 we can convert a desired Cartesian trajectory into a joint-angle trajectory and ask the humanoid robot to track it, or simply we can ask the robot to perform a reaching movement for a single desired end-effector position without falling over.

The open parameters N, the number of basis functions, and σ^2 (the common variance of all the basis functions) were determined using cross-validation. The parameters $\boldsymbol{\mu}_i$ (the means of the basis functions) were determined by taking N even samples from the

trajectory generated by the subject. For cross-validation, we prepared a Cartesian desired trajectory that was not part of the recorded data set, and converted it into a joint trajectory with the current candidate values of (N,σ^2). The joint trajectory was simulated on a kinematic model of the humanoid robot to produce an end-effector trajectory. To select the best parameters from a fixed set of candidate (N,σ^2) pairs, we used a deviation measure of the resultant trajectory from the desired trajectory defined as the mean value of the sum of the absolute error in all joints

$$deviation = \frac{1}{m}\sum_{i=1}^{m} e_i, \tag{7}$$

where m is the total number of data points of the trajectory used for cross-validation and e_i is the absolute error in all K joints defined as

$$e_i = \sum_{j=1}^{K} |q^j_{pred} - q^j_{des}|. \tag{8}$$

The values determined for N and σ using this scheme were 20 and 60, respectively.

5 Statically stable motion generation on the humanoid robot

The use of radial basis functions to form a nonlinear mapping from end-effector positions (**X**) to joint angles (**Q**) yielded a robust inverse kinematics mapping that preserved the (static) balance of the robot. We empirically observed that the robot could reach the space in the convex hull of the end effector positions obtained in the data collection (refer to Figure 3). Because during data collection, human movement—and hence the robot movement—was slow, the data collected essentially ignored the possible dynamics effects of reaching. Even though the inverse kinematics solution we obtained may be correct, a fast reach to that target, or continuous tracking of a Cartesian trajectory may make the robot fall. Therefore, it was not clear to what extent the obtained inverse kinematics would generalize to cases where the dynamics is not negligible and/or continuous tracking is required. To test this, we prepared an elliptical Cartesian trajectory that the robot was required to track using the obtained inverse kinematics mapping, which may be run at different speeds. The trajectory was defined via the parametric curve

$$x(t_i) = \begin{bmatrix} cos\alpha & -sin\alpha \\ sin\alpha & cos\alpha \end{bmatrix} \begin{bmatrix} 3\,sin(t_i) - 9 \\ 15\,cos(t_i) + 26 \end{bmatrix}, \tag{9}$$

where $t_i = 0,\ 0.05,\ 0.1,\ldots,2\ \pi$ and $\alpha = \pi/18$. For the robot execution, the desired trajectory $\boldsymbol{x}(t_i)$ was mapped to $\boldsymbol{q}_{pred}(t_i)$ using Equation 6. Afterwards a smooth continuous function $\boldsymbol{q}(t)$ was obtained by cubic spline interpolation of the data points $(t_i,\boldsymbol{q}_{pred})$ so that robot stability could be tested at different speeds of trajectory tracking. Figure 5 shows a sequence of video frames illustrating the autonomous tracking of this trajectory at the nominal speed (chosen to have the relative static stability level within the limits of data collection, see Figure 4). The tracking performance with this speed is shown in Figure 6 where the desired end-effector trajectory and the actual trajectory achieved by the robot through the use of Equation 6 are shown. The robot tracked the elliptical trajectory faithfully as the trajectory deviation defined by Equations 7 and 8 was 0.7°. The light wiggly curve also shown in Figure 6 shows the end-effector trajectory that was generated by the human demonstrator in the first phase, and which was subsequently used to determine **W**, defining the nonlinear relationship between the joint angles and the end-effector position.

In principle, the paradigm was used to obtain *statically stable* reaching; so, when the robot was asked to perform a trajectory at speeds significantly higher than previously demonstrated, the inertial dynamics of the robot must have become significant. The generalization of the skill over different inertial dynamics is illustrated in Figure 7 where the upper plot shows the stability measure when the elliptical trajectory tracking was performed at 0.1 Hz (i.e., one complete loop in 10 s). When the motion was performed at twice this speed, the robot became unstable as can be seen in the lower panel of Figure 7. The robot could still track the desired trajectory without falling over, but just barely. When the speed was increased any further the robot became unstable and fell over. So, we can conclude that with this particular robotic system, the statically balanced reaching could generalize to speeds twice that of the nominal speed. This level of generalization was somewhat surprising as the subjects were not required to learn about dynamics, as the application of the HL-RSS paradigm involved only static stability requirements.

6 Dynamic posture control through HL-RSS

Up to now, the application of the HL-RSS paradigm was limited to cases where human subjects did not need to handle the dynamics (i.e., they could choose to move slowly). This section describes a new experiment, separate from the statically reaching task, where subjects are involved in a true dynamics task within the framework of this paradigm. In this new experiment, we

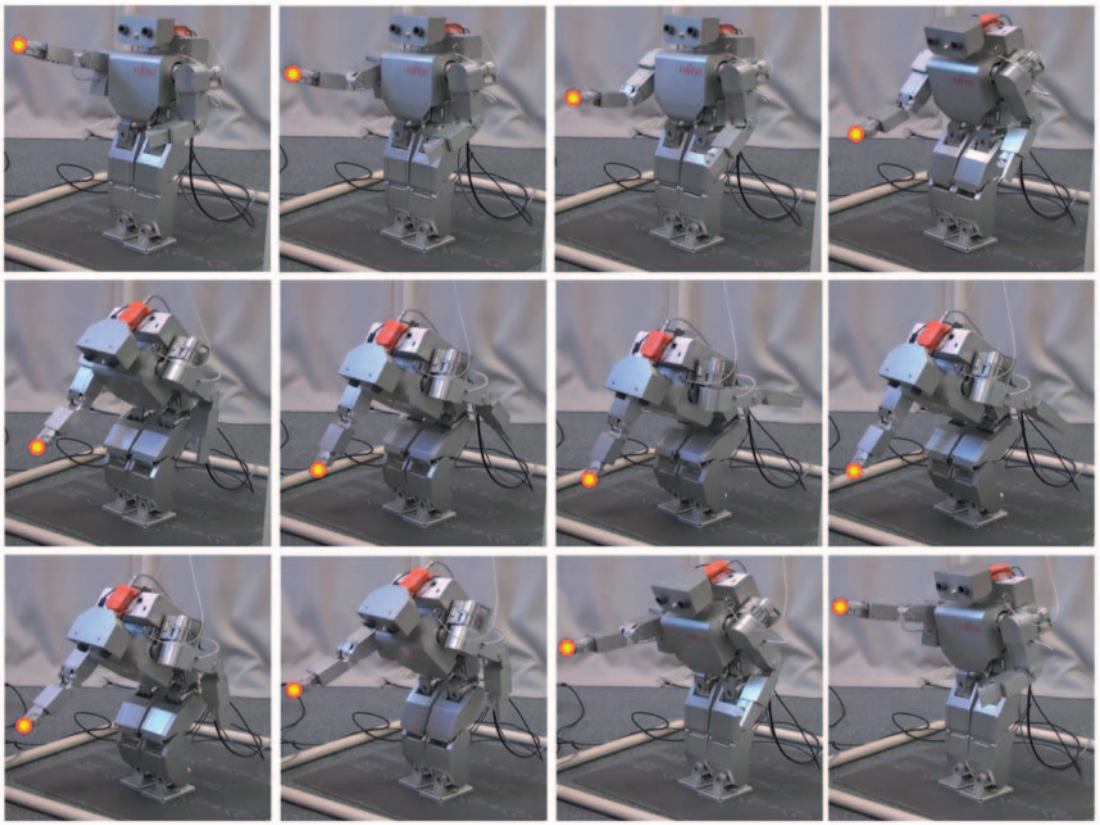

Figure 5. Video frames showing the statically stable reaching motion obtained by the humanoid robot.

tested whether a human could learn to generate the motor commands required to maintain the vertical balance of an inverted pendulum modeling the dynamics of a humanoid robot's center of mass. We subsequently examined the comparative effects of visual and proprioceptive-vestibular feedback on human learning. For visual feedback we utilized a computer screen. For proprioceptive-vestibular feedback, however, we constructed a platform with a parallel 6 degree-of-freedom Stewart-type mechanical structure, capable of carrying a human subject. The parallel platform is shown in Figure 8. The diameter of the top plate of the platform is 0.7 m and it is actuated by six linear electric servo drives.

A successful HL-RSS design provides a setup in which a human's natural response to a well-accustomed control task (e.g., controlling one's own posture) also generates robot motor commands that help the robot to achieve the target control task (e.g., robot stays upright). In the case of visual feedback, there is no physical connection between the feedback and the motion constituting the human's response: it is an arbitrary mapping learned through practice. However, in the proprioceptive-vestibular feedback case, and in particular with the platform, the robot control task (keeping the robot upright) and the human's natural motor control (maintaining their posture) can be aligned in a synergistic way so that the movements the human performs to maintain body balance also generate commands which maintain the robot's balance. This is the primary motivation for investigating the comparative effect of visual and proprioceptive-vestibular feedback empirically, thus calling for a proprioceptive-vestibular feedback system.

In the experimental sessions, subjects stood on the platform while their motion was captured in real time using an active marker-based motion capture system (3D Investigator, Northern Digital Inc.). The angle of the human's hip joint was calculated from the positions of three markers placed on the ankle, hip, and shoulder joints. For each frame of motion capture data, the human posture was used to provide the input torque for a simulated inverted pendulum that modeled a humanoid robot with a mass of 70 kg, distributed around a center of mass at a height of 0.98 m (which roughly corresponds to a 70 kg, 175 cm tall humanoid robot). The magnitude of the torque was calculated by multiplying the hip angle by 50 N·m/deg. This multiplying factor was calculated based on the average maximal extension of the hip joint in adults (Boone & Azen, 1979) and the torque needed to recover the pendulum from its limit of 18°.

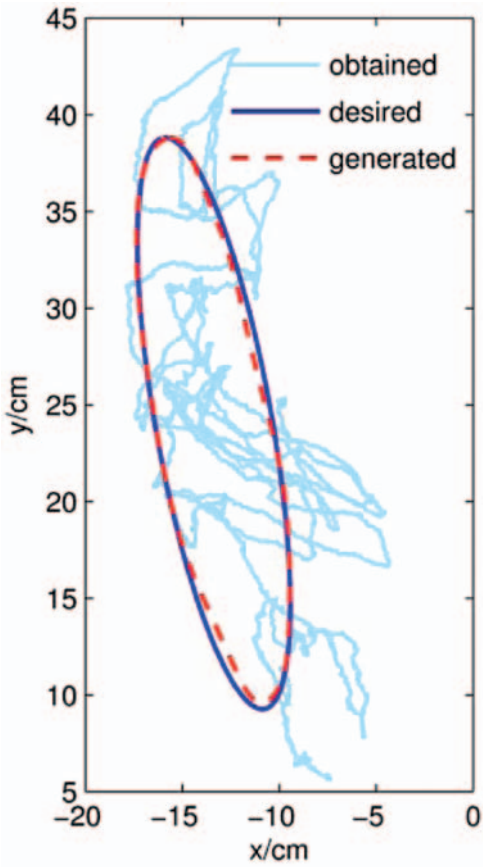

Figure 6. The end-effector trajectory generated by the demonstrator (light wiggly curve), with the desired end-effector trajectory (solid line), and the generated end-effector trajectory (dashed line) obtained by executing the deduced joint-angle trajectory.

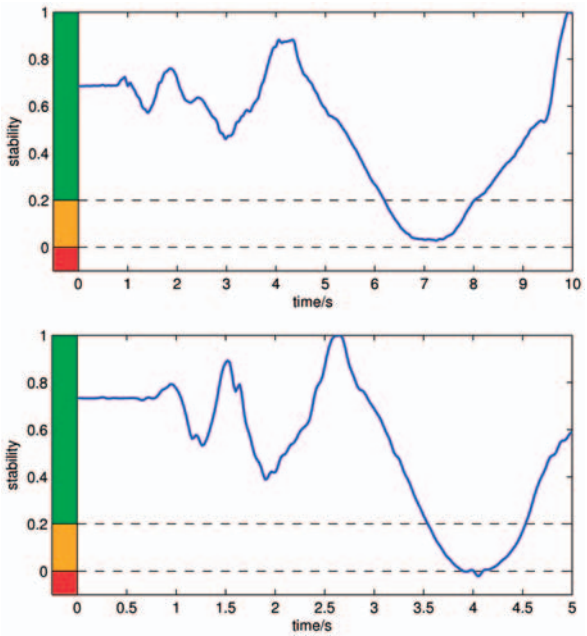

Figure 7. The stability of the humanoid robot when the elliptical trajectory shown in Figure 6 was performed at different speeds. When the robot moved quickly, the inertial effects were no longer negligible, as illustrated by the minimum occurring below 0 after around 4 s in the lower plot.

Figure 8. Stewart-type parallel platform with 6 degrees of freedom, capable of carrying a human subject.

Specifically, the 50 N·m/deg enables an average male adult at the maximal hip extension to generate roughly twice the torque needed to maintain the inverted pendulum used in our experiment at its maximal angle of 18°. The task of the human was to preserve the upright pose of the inverted pendulum by moving the body. Thus, during the experiment, the human posture continuously generated a torque to drive the inverted pendulum. In return, the motion of the pendulum, namely the angle of the joint, was fed back to the human in real time through either visual rendering of an animated humanoid robot or by rotating the platform on which the human was standing. The experimental setup is shown in Figure 9.

Eight subjects participated in the experiments which spanned two consecutive days. On the first day, the task of the subjects was to learn to keep the upright pose of the pendulum without any external perturbations. On the second day, the pendulum was randomly perturbed via short torque pulses which disturbed the posture of the pendulum, requiring the subjects to perform immediate corrective movements. Half of the subjects first performed the experiment with visual feedback, and

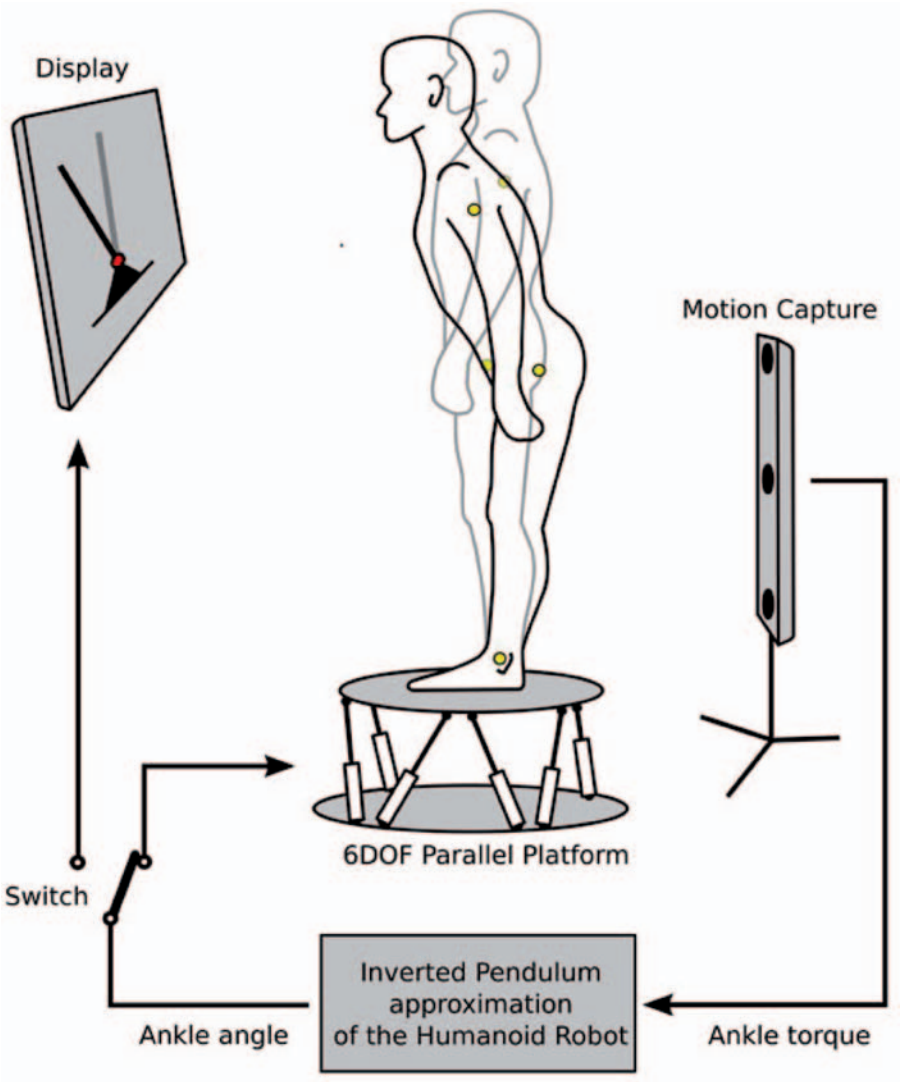

Figure 9. Experimental setup.

the other half of the subjects first performed the experiment with the proprioceptive-vestibular feedback. This was so for both first and second day experiments. Then, they took a 2-hr break. After the break, the feedback modality was switched so that the group of subjects, who performed the pre-break experiment with the visual feedback, then performed the experiment with the proprioceptive-vestibular feedback. The other half of the subjects, who performed the pre-break experiment with the proprioceptive-vestibular feedback, then performed the experiment with the visual feedback. To provide the proprioceptive-vestibular feedback to the subject, the platform tilt angle was connected to the simulated pendulum angle through a high gain PD servo. The control frequency of the parallel platform was 500 Hz. In the visual feedback case the subjects stood on the platform, but the platform never moved.

7 Analysis of human control of the simulated inverted pendulum

To test whether the human is able to balance the pendulum with the size and mass of the full size humanoid robot utilizing either vision or proprioceptive-vestibular feedback, the subjects first had 15 minutes available to learn to maintain the stable upright posture of the pendulum with both feedback modalities, each implemented separately. All subjects reported feeling comfortable performing the task with either vision or proprioceptive-vestibular feedbacks well before the perturbation trials. Having confirmed that both modalities were sufficient for unperturbed control, we focused on the more challenging case of perturbed trials where the balance of the pendulum was assailed by short random torque pulses.

We first examine the trajectory of the pendulum after the injection of torque impulses. The top two diagrams shown in Figure 10 show several example trajectories after a perturbation (shown in green). The left diagram shows the pendulum angle trajectories (red lines) when the subject controlled the pendulum through visual feedback. The right diagram shows the pendulum angle trajectories (blue lines) when the subjects received the pendulum state through proprioceptive-vestibular feedback. Inspecting the left and right trajectories, it can be seen that the time needed to return the pendulum to the upright position (angle $= 0^\circ$) is shorter with the proprioceptive-vestibular feedback than with the visual feedback.

In order to investigate the responses of the human subjects in the frequency domain, we transformed the torque supplied by the subjects to the pendulum into the frequency domain. The bottom two diagrams show the power spectrum density of the torque generated by the human during the experiment. The green peaks correspond to the frequency distribution of the perturbation pulses while the red and blue lines show the frequency distributions of the torque generated by the human with the visual and proprioceptive-vestibular feedback. In both cases the main peaks of the power spectrum density correspond well with the frequency distribution of the perturbation pulses. However, it can be seen that the amplitudes in the case when the visual feedback was used are significantly higher than the amplitudes of the power spectrum density when proprioceptive-vestibular feedback was used. Moreover, in the visual feedback case there are some components of the torque that lie between the main peaks, which suggests that the response of the human was not as closely synchronized with the perturbation signal as in the case of the proprioceptive-vestibular

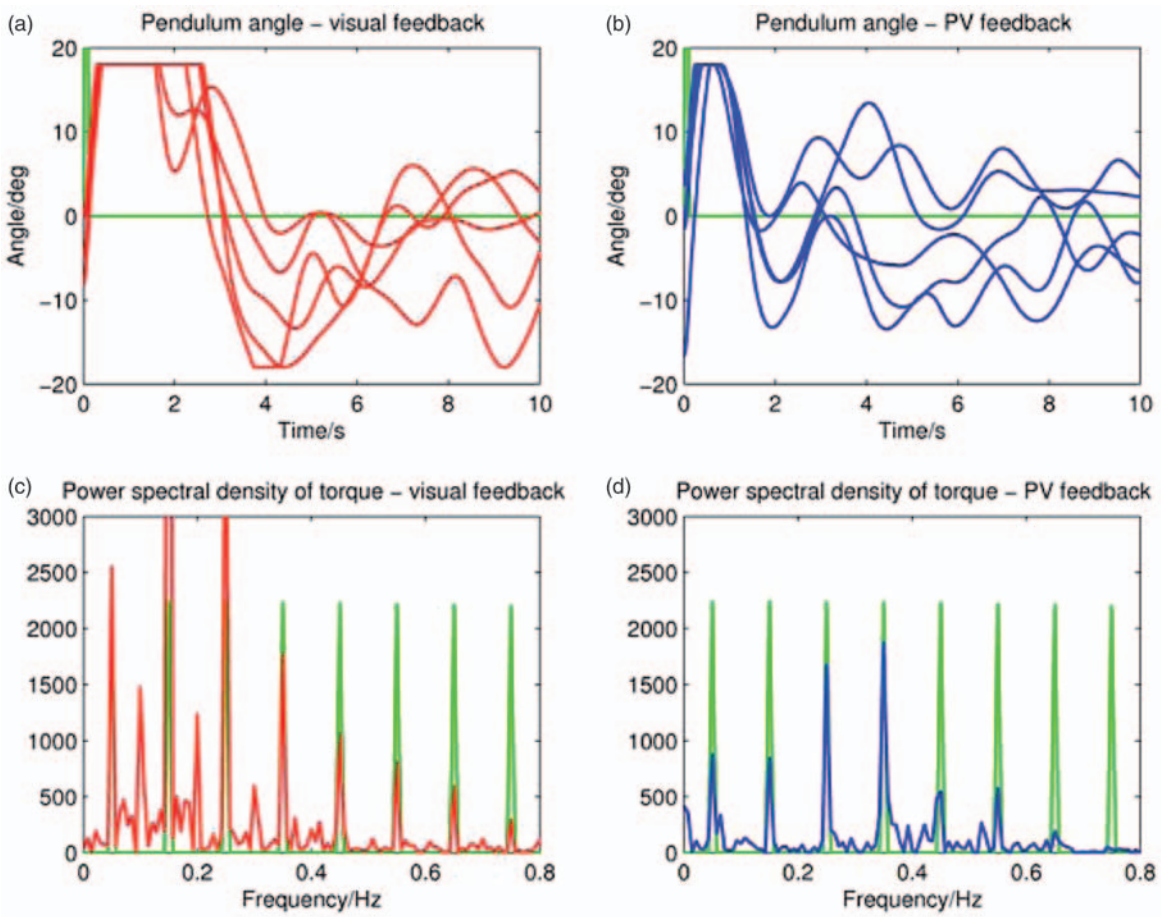

Figure 10. Sample pendulum-angle trajectories after perturbations (upper plots) and power spectrum density of torque (lower plots) for visual (left) and proprioceptive-vestibular (PV) (right) feedback setups. The time needed for the pendulum to return to the upright position is shorter, and the power injected into the system is lower with the proprioceptive-vestibular feedback. Note that the thresholding of the pendulum angle seen in the first few seconds is due to the limiting of the platform (and hence pendulum) angle to $[-18^\circ, 18^\circ]$ for the sake of safety.

feedback. This can be explained by the fact that the motion of the human was more tightly connected with the motion of the pendulum through the presentation of proprioceptive-vestibular feedback (rotation of the platform) that directly affected the human's own balance. This does not directly change the torque supplied, but the reflex motions of the subject for maintaining their own posture indirectly induce a torque change in the direction of pendulum stability. When visual feedback is provided instead, the human experiences the balance state of the pendulum indirectly (by observing it on the display), and their reactions are more dependent on the cognitive mapping they establish in order to balance the pendulum.

We defined two metrics with which to quantitatively compare the visual feedback and proprioceptive-vestibular feedback cases of inverted pendulum control. The first is the reaction time of the human with respect to a perturbation, t_s. This is defined as the time from the onset of the perturbation pulse until the next instant the pendulum crosses the upright position. The second is the virtual energy expenditure of the human, E, required to keep the pendulum upright. This is approximated by the integral of the absolute value of the pendulum angle during the 10 s after a perturbation,

$$E = \int_0^{10s} |\theta(t)| \, \mathrm{d}t, \tag{10}$$

where $\theta(t)$ is the pendulum angle at time t. An independent samples t test was conducted to compare the reaction times and energy expenditures for the visual feedback and proprioceptive-vestibular feedback cases. For the reaction time, there was a significant difference in the scores for visual feedback ($M = 2.40$ s, $SD = 0.84$ s) and for proprioceptive-vestibular feedback ($M = 1.62$ s, $SD = 0.62$ s) conditions; $t(54) = 6.89$, $p = 6.78\text{E}{-9}$. Also for the energy measure, there was a significant difference in the scores for visual feedback ($M = 1.77$ s, $SD = 0.38$ s) and proprioceptive-vestibular feedback ($M = 1.48$ s, $SD = 0.38$ s) conditions; $t(54) = 4.37$, $p = 5.46\text{E}{-5}$. These results indicate that during real-time control with proprioceptive-vestibular feedback there is a significant decrease in both the reaction time and the virtual energy expenditure (which is closely related to, and can be considered a metric for, real energy consumption).

8 Discussion

The first experiment exemplifies the application of the HL-RSS paradigm for a humanoid robot skill, namely statically stable reaching. The key observation there is that although the skill itself is computationally nontrivial (though solvable), even a little child, let alone a nonexpert, can equip the robot with this skill through this paradigm. If one hopes for a future with robots helping us in our daily lives, then we think this is the right direction to pursue. If the experts are to program every tiny skill for robots then this dream may stay as a dream for long years to come. We envision an era where the robot skills are generated by robot owners and shared within the community in the spirit of shareware software distribution over the Internet. Then, not only these custom skills, but also primary modes of operation of these robots would be obtained in this fashion. Robot companies would hire people as robot trainers to build up basic skills of their robots, such as walking. We do not think classical robotics would become obsolete, on the contrary it would be essential in providing the basic firmware for robots on top of which new behaviors could be built. When we imagine a computer, it is automatically the hardware plus the operating system running on it. In the future, this might be much the same for the robots too; the classical robotics and control techniques would be at work for making up the operating system of a robot.

Of course there are major issues to be tackled, such as safety, and other issues that are important for allowing nonexperts to teach robots intuitively, such as joy of teaching.[1] For example in the first experiment we carried out, the subjects had the feeling of playing a game, and felt a sense of achievement when they could complete a session without violating the rules of the game, that is, keeping the cursor (center of mass of the robot projected on the ground) in the displayed polygon (the support polygon of the robot). The subjects could learn to follow the rule of the game in less than 20 min. This fast learning was probably due to the static nature of the task. For example, if we asked subjects to punch targets in space while keeping the cursor in the polygon, presumably they would need longer training. The intuitiveness is important for fast learning as we have seen in our previous work where the HL-RSS paradigm was applied to a manipulation task. The task there was to swap the position of two balls without dropping them using a robotic hand which was controlled in real time by a human subject (Oztop et al., 2006). This task was not intuitive in the sense that when one performs the ball swapping in one's hand in natural conditions, vision is not used at all; the task is mainly guided by the tactile sensory system. Although the subject could learn to control the robot to swap the balls, the process was rather effortful and took a week of training (1–2 hr/day) (Oztop et al., 2006). So, we have reasons to think that compatibility of the feedback modality to the control task is important for intuitiveness, speed of learning, and fun factor.

Our second experiment aimed to test human performance issues based on the feedback modality. In addition, we also wanted to verify that humans can learn to control a robot when dynamics of the robot cannot be ignored. For this, the most straightforward choice was an inverted pendulum approximating a full-sized humanoid robot. The human movements supplied the torque to the inverted pendulum. The task was to keep the pendulum straight up. The feedback modality we chose to compare was a third person view of the robot versus the "sense of balance." Our prediction was that sense of balance would be more appropriate for controlling the upright posture of the pendulum. To our surprise, in both cases humans could quickly learn to keep the robot upright, even in the face of perturbations. This was probably due to the slow time constant of an inverted pendulum approximating a 70 kg, 175 cm tall humanoid robot. Nevertheless, we could see some differences, such as early recovery of the perturbation and less energy consumption with the sense of balance. However, there are issues we cannot resolve with the current setup. When the subjects are given feedback by the platform, we cannot single out the sensory system that is the determinant for these changes. The sense of balance is conveyed through multiple sensory systems: the proprioceptive, vestibular, and visual systems. Because the latency of proprioceptive input is shorter than visual and vestibular inputs and tightly linked with the stretch reflex (Fitzpatrick, Gorman, Burke, & Gandevia, 1992), the difference in performance can be attributed to proprioception. Another possibility, which is not exclusive, is that the convergence of multiple sensory signals may make it possible to obtain a more accurate sense of balance. One point we can safely make is that the platform stimulates the three sensory systems in a coherent way. Therefore, we may speculate that the subjects engage their natural responses for posture control allowing them to bring back the pendulum to its desired upright state intuitively with the platform.

9 Conclusion

A major goal of robotic imitation of a demonstrated motion is to remove the burden of robot programming from experts by having nonexperts teach robots. The most basic method to transfer a given motion from a demonstrator to a robot is to directly copy the motor commands of the demonstrator to the robot (Atkeson et al., 2000). Although this approach is very efficient for certain open-loop tasks, it is generally not possible to adopt it: either the motor commands may not be available to the robot, or the differences between the demonstrator's body and the robot may be so large that a direct transfer of motor commands is not possible. We have described a paradigm based on the application of human learning for robot skill synthesis (HL-RSS), which differs because the correct motor commands for the robot are produced by the human controlling the robot in real time. The price for this convenience is the training the human must undergo in order to control the robot to successfully complete a given task. In short, instead of expert robot programming, the HL-RSS paradigm relies on human sensorimotor learning capabilities to produce appropriate motor commands on the robot, which can subsequently be used to derive autonomous controllers through machine learning methods.

The main contribution of the current study is the introduction of the HL-RSS paradigm and its application to full body humanoid robot skill synthesis and an analysis of the feedback interface portion of the paradigm. In particular, we showed how the skill of statically stable reaching could be obtained for a humanoid robot using the HL-RSS paradigm. We demonstrated that the paradigm is applicable to tasks involving nonnegligible dynamics by having human subjects control the upright posture of an inverted pendulum modeling a full-sized humanoid robot. Finally, we investigated the effect of the feedback modality used to present the robot state to the human during real-time robot control. The results confirmed that for an ideal HL-RSS design, the setup should induce natural responses from the human that would help the robot to achieve the desired task. For posture control, this is achieved by having a Stewart platform that not only feeds the state of the robot back to the human but also *compels* the human to move so as to maintain self postural stability, which, in turn, generates motor signals that help the robot maintain its balance.

Note

1. Here the phrase "to teach robots" actually refers to the sensorimotor learning to enable a person to control the robot for performing a target task with it, such as operating a robot to make it paint a wall.

Funding

This work was supported by the Japan Society for Promotion of Science and the Slovenian Ministry of Higher Education, Science and Technology (J. Babič); and the Global COE Program Center of Human-Friendly Robotics Based on Cognitive Neuroscience at the Ministry of Education, Culture, Sports, Science and Technology, Japan (E. Oztop).

References

Atkeson, C., Hale, J., Pollick, F., Riley, M., Kotosaka, S., Schaal, S. et al. (2000). Using humanoid robots to study human behavior. *IEEE Intelligent Systems*, *15*, 45–56.

Bentivegna, D., Atkeson, C., & Cheng, G. (2004). Learning tasks from observation and practice. *Robotics and Autonomous Systems*, *47*, 163–169.

Billard, A., & Hayes, G. (1999). Drama, a connectionist architecture for control and learning in autonomous robots. *Adaptive Behavior*, *7*, 35–63.

Billard, A., & Siegwart, R. (2004). Robot learning from demonstration. *Robotics and Autonomous Systems*, *47*, 65–67.

Boone, D., & Azen, S. (1979). Normal range of motion of joints in male subjects. *The Journal of Bone and Joint Surgery*, *61*, 756–759.

Breazeal, C., & Scassellati, B. (2002). Robots that imitate humans. *Trends in Cognitive Sciences*, *6*, 481–487.

Chen, J., Haas, E., & Barnes, M. (2007). Human performance issues and user interface design for teleoperated robots. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, *37*, 1231–1245.

Fitzpatrick, R. C., Gorman, R. B., Burke, D., & Gandevia, S. C. (1992). Postural proprioceptive reflexes in standing human subjects: bandwidth of response and transmission characteristics. *Journal of Physiology*, *458*, 69–83.

Goldberg, G., & Hagmann, S. (1998). Tool use and mechanical problem solving in apraxia. *Neuropsychology*, *36*, 581–589.

Imamizu, H. (2010). Mechanisms of human sensorimotor-learning and their implications for brain communication. *IEICE Transactions on Communications* (Vol E91.B, pp. 2102–2108).

Ito, M., & Tani, J. (2004). On-line imitative interaction with a humanoid robot using a dynamic neural network model of a mirror system. *Adaptive Behavior*, *12*, 93–115.

Kushida, D., Nakamura, M., Goto, S., & Kyura, N. (2001). Human direct teaching of industrial articulated robot arms based on force-free control. *Artificial Life and Robotics*, *5*, 26–32.

Oztop, E., Lin, L.-H., Kawato, M., & Cheng, G. (2006). Dexterous skills transfer by extending human body schema to a robotic hand. In *Paper presented at the IEEE-RAS International Conference on Humanoid Robots*. Genova, Italy.

Oztop, E., Lin, L.-H., Kawato, M., & Cheng, G. (2007). Extensive human training for robot skill synthesis: Validation on a robotic hand. In *Paper presented at the IEEE International Conference on Robotics and Automation*. Roma, Italy.

Park, J., & Sandberg, I. W. (1991). Universal approximation using radial-basis-function networks. *Neural Computation*, *3*(2), 246–257.

Schaal, S. (1999). Is imitation learning the route to humanoid robots? *Trends in Cognitive Sciences*, *3*, 233–242.

Shadmehr, R., & Wise, S. (2005). *Computational neurobiology of reaching and pointing: A foundation for motor learning*. Cambridge, MA: MIT Press.

Shidara, M., Kawano, K., Gomi, H., & Kawato, M. (1993). Inverse-dynamics model eye movement control by Purkinje cells in the cerebellum. *Nature*, *365*, 50–52.

Steinfeld, A., Fong, T., Kaber, D., Lewis, M., Scholtz, J., Schultz, A. et al. (2006). Common metrics for human-robot interaction. In *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*. Salt Lake City, USA (pp. 33–40).

Ude, A., Gams, A., Asfour, T., & Morimoto, J. (2010). Task-specific generalization of discrete and periodic dynamic movement primitives. *IEEE Transactions on Robotics*, *26*(5), 800–815.

Wolpert, D., & Ghahramani, Z. (2000). Computational principles of movement neuroscience. *Nature Neuroscience*, *3 Suppl*, 1212–1217.

Wolpert, D., Ghahramani, Z., & Flanagan, J. (2001). Perspectives and problems in motor learning. *Trends in Cognitive Sciences*, *5*(11), 487–494.

About the Authors

Jan Babič is a research fellow at the Jožef Stefan Institute and a senior researcher in the Faculty of Electrical Engineering, University of Ljubljana, Slovenia. He received his Ph.D. from the Faculty of Electrical Engineering, University of Ljubljana in 2004. During 2006 and 2007 he was a visiting researcher at the ATR Computational Neuroscience Laboratories in Japan. His current research is concerned with understanding how the human brain controls movement of the arms and legs, and with the design of biologically plausible robot controllers that achieve similar robustness and adaptation to the changing environment as found in humans.

Joshua G. Hale received a B.A. (Hons. 1st) in computation (Oxford University, 1997), M.S. (Dist.) in computer science (Edinburgh University, 1998), M.A. in computation (Oxford University, 2002), and Ph.D. on biomimetic motion synthesis (Glasgow University, 2003). He worked as a research engineer in the Oxford University Hardware Compilation Group, research assistant at the Glasgow University Computer Vision and Graphics Laboratory, researcher in the ATR Humanoid Robotics and Computational Neuroscience Laboratory, and is currently a senior researcher at Cyberdyne Inc., where he works with interactive devices, dynamic simulation, and computer graphics. He enjoys gardening, singing, Thai boxing, and baking bread. *E-mail*: joshua_hale@cyberdyne.jp

Erhan Oztop earned his Ph.D. at the University of Southern California in 2002. Right after this, he joined ATR Computational Neuroscience Laboratories as a researcher, and later became a senior researcher (2002–2011). He was also involved in JST, ICORP Computational Brain Project as a researcher and later a group leader (2004–2008). Currently, he is with NICT, Advanced ICT Research Institute, Brain ICT Laboratory, Japan and is leading the Communication and Cognitive Cybernetics group at ATR, Cognitive Mechanisms Laboratories. He also holds a visiting associate professor position at Osaka University. *E-mail*: erhan@atr.jp