Learning to Maintain Upright Posture: What Can Be Learned Using Adaptive Neural Networks Models?

N. Alberto Borghese[1], Andrea Calvi[1,2]
[1]Laboratory of Human Motion Analysis and Virtual Reality (MAVR), Department of Computer Science, University of Milano
[2]Department of Bioengineering, Politecnico of Milano

Human upright posture is an unstable position: Continuous activation of postural muscles is required to avoid falling down. This is the output of a complex control system that monitors a very large number of inputs, related to the orientation of the body segments, to produce an adequate output as muscle activation. Complexity arises because of the very large number of correlated inputs and outputs: The finite contraction and release time of muscles and the neural control loop delays make the problem even more difficult. Nevertheless, upright posture is a capability that is learned in the first year of life. Here, the learning process is investigated by using a neural network model for the controller and the reinforcement learning paradigm. To this end, after creating a mechanically realistic digital human body, a feedback postural controller is defined, which outputs a set of joint torques as a function of orientation and rotation speed of the body segments. The controller is made up of a neural network, whose "synaptic weights" are determined through trial-and-error (failure in maintaining upright posture) by using a reinforcement learning strategy. No desired control action is specified nor particular structure given to the controller. The results show that the anatomical arrangement of the skeleton is sufficient to shape a postural control, robust against torque perturbations and noise, and flexible enough to adapt to changes in the body model in a short time. Moreover, the learned kinematics closely resembles the data reported in the literature; it emerges from the interaction with the environment, only through trial-and-error. Overall, the results suggest that anatomical arrangement of the body segments may play a major role in shaping human motor control.

Keywords   reinforcement learning · posture · neural networks · learning with a critic

1   Introduction

Upright human posture is an unstable position. It can be compared to the vertical position of an inverted pendulum: Small external perturbations do produce falling down. To avoid this, postural muscles are continuously activated to compensate for perturbations. Their activation is the output of a complex control system, whose complexity arises from several reasons, both structural and functional. Because of the articulated nature of the skeleton, the torque produced by a single muscle affects all the body segments: In control theory terminology, this is a multi-input/multi-output system, with a very large control space (e.g., the leg has 56 major muscle groups); due to its nonlinearity it cannot be dealt with using classical control

Correspondence to: N. A. Borghese, Laboratory of Human Motion Analysis and Virtual Reality (MAVR), Department of Computer Science, University of Milan, Via Comelico 39, 20135 Milan, Italy. *E-mail*: borghese@dsi.unimi.it, http://homes.dsi.unimi.it/~borghese/; *Tel.*: +39–02–50316325 + *Fax:* +39–02–50316373

algorithms (Astrom & Wittenmark, 1989). From the functional point of view, the control system has to take into account the finite power and bandwidth of muscles (Kashima & Isurugi, 1998) and the time delay produced by the neural control loops (Ghez, 1991). Moreover, it has to take into account the geometrical arrangement of the body segments and the mass distribution inside the segments (de Leva, 1996). Nevertheless, maintaining upright posture is a capability that is learned in the first year of life and maintained with no conscious mental effort throughout life (Forssberg, 1999).

Several experimental studies have tried to investigate the strategies used by the central nervous system (CNS) to accomplish this task (Massion, 1994; Lacquaniti 1997; Winter, Patla, Prince, Ishac, & Gielo-Perzak, 1998; Forssberg, 1999). According to one point of view (Winter et al., 1998), CNS sets an adequate stiffness value for the postural muscles, mainly the ankle ones: When the body sways, its vertical position is restored by visco-elastic torques elicited by muscle stretch. This pure passive behavior has been questioned by Morasso and Schieppati (1999) who showed that active mechanisms are indeed required to stabilize body sway. According to a different point of view, the CNS tries to keep the center of mass inside the support base (the feet). The center of the foot line on the floor acts as an attractor of the body: Fluctuations around this point have been shown to obey chaotic attractor laws (Collins & De Luca, 1994). Although this is a good description of postural kinematics, it does not explain how upright posture is maintained.[1] Further experimental investigation has been aimed at understanding this. The results support the original proposal of Nicolaj Bernstein (1967) of the CNS controlling synergies. Synergies are ensembles of elements that act as a single entity, a unicum, for a given goal. The group of Lacquaniti has stressed the importance of kinematic synergies (Lacquaniti & Maioli, 1995; Lacquaniti, 1997). In this view, the orientation of the body segments both in humans and in cats is the variable controlled by the CNS to maintain upright posture. Moreover, the observation of covariations among the different segments has led some researchers to postulate the presence of kinematic synergies, which would simplify the control task even more (e.g., Mah, Hulliger, Lee, & O'Callaghan, 1994). This stress on the kinematics follows from the postulate of a body scheme (Gurfinkel, Levik, Popov, Smetanin, & Shilkov, 1988)

that persists even when external conditions change (e.g., in microgravity, Mounchino, Cincera, Fabre, Assainte, Amblard, Pedotti, & Massion, 1996).

A further question can be posed: Are these synergies hardwired in our genetic code (ontogenetically developed), or are they the result of the interaction of our body (as a dynamical model), with the environment? Are there genetic synergies that shape the postural control, or is it the postural control that shapes the synergies? In this article the possibility that postural synergies are shaped by the interaction of the human body with the environment with no a priori constraint is investigated. To this aim, learning upright posture has been studied using a neural network model for the controller (Miller, Sutton, & Werbos, 1990) and the reinforcement learning paradigm (Kaelbling et al., 1996; Doya, 2001). No information is given to the system but failure in maintaining upright posture. Results show that the controller does learn to maintain upright posture and that the resulting time course of the kinematics resembles the one described in the literature.

The article is organized as follows. The controller and the dynamical model of the muscles and the body are introduced in Section 2. Section 3 describes the reinforcement learning procedure and Section 4 the simulations. Results are described in Section 5, where results of experiments with torque disturbances and modification of the body model are also reported. The results are discussed in Section 6 and a brief conclusion with possible further developments is drawn in Section 7.

2 Methods

The schematization of the postural control task is shown in Figure 1. The dynamical model of the human body is subjected to a set of torques, the muscular torques $\mathbf{T}(t)$, which are obtained by the transformation, through a dynamical model of the muscles, of a set of neural signals, $\mathbf{n}(t)$. These neural signals are produced by the controller, as a function of the body segments' orientation, $\mathbf{a}(t)$, and rotation speed, $\dot{\mathbf{a}}(t)$. The $\mathbf{n}(t)$ are then transformed into torques, $\mathbf{T}(t)$, through a dynamical model of the muscles. The controller is a parametric adaptive model whose parameters have to be adequately set to maintain the body upright. No information is available about good values for the parameters; the only information available is the failure of the
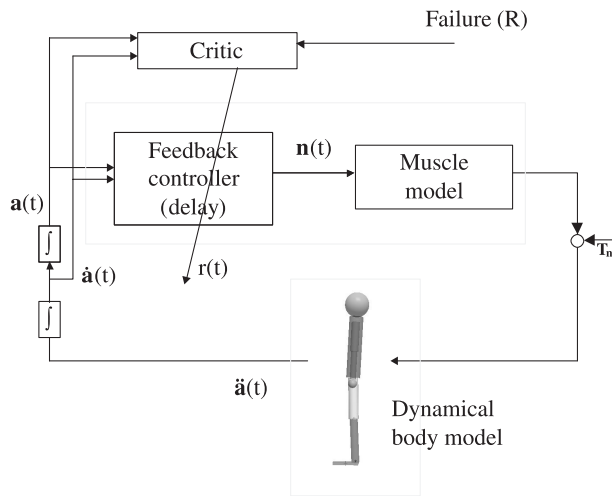
Figure 1 The model. The feedback controller produces a set of neural signals, **n**(t), as function orientation and rotation speed of the body segments. The **n**(t) are transformed into muscle torques, **T**(t), according to a dynamical muscle model. Torques, in turn, produce a change in the kinematics of the body model. The dynamical model of the body is composed of nine macro-segments: the HAT (head, arms, and trunk) and the two legs (upper and lower leg, and foot). The controller action depends on a set of parameters that are tuned by a higher module: the critic. This monitors the time sequence of the orientation and rotation speed of the body segment and produces an internal reinforcement signal, $r(t)$, which is used to adapt the parameters of the controller. **T_n** represents the torque used as a a perturbation in the experiments described in Section 5.1.

task. This is an instantaneous bit of information that has to be transformed into a signal continuous in time, suitable to adapt the controller parameters. This is the job of a second module called the critic. The different modules are examined in more detail below.

2.1 The Dynamical Model of the Human Body

To achieve a high reliability in the simulation, an accurate dynamical model of the skeleton and of the musculo-skeletal system has been implemented. The human skeleton has been modeled as an articulated chain, composed of a set of rigid segments connected by hinges (Pedotti, 1977; Winter, 1990; de Leva, 1996). Although human joints exhibit complex motions and cannot be defined as simple spherical joints, their representation as hinges is usually accepted in simulations related to whole-body motion control (Winter, 1990; Anderson & Pandy, 2001). The hinges are aligned along the vertical so that gravity is counterbalanced by

the vinculum reaction generated by the support of the feet and the body is in equilibrium. This is an unstable equilibrium position.

To describe the dynamics of the body model, the length and the mechanical properties of each segment (mass, center of mass position, and inertial moment) are required: These quantities were derived from the biomechanics literature (Zatsiorsky, Raitsin, Seluyanov, Aruin, & Prilutzky, 1993, de Leva, 1996; cf. Table 1).

On one side, this is a large number of parameters that makes mechanical simulations slow, while on the other side, the relative motion of some of these segments is not relevant to problems related to posture. In this work, the upper trunk, the arms, and the head are lumped into a single segment, the HAT (head, arms, and trunk), as it is referred to in the literature (e.g., Winter, 1990; Anderson & Pandy, 2001). The mechanical parameters of the HAT have been computed from the parameters of its constituent segments. As a result, the body model considered here in the simulations is constituted of seven segments: the HAT and the two legs composed of three segments each: the thigh, the leg, and the foot. The dynamics of the body model can therefore be represented as

$$\ddot{\mathbf{a}}(t) = \mathbf{k}(\mathbf{T}(t), \mathbf{a}(t), \dot{\mathbf{a}}(t) \mid \mathbf{z}) \tag{1}$$

where $\mathbf{a}(t)$, $\dot{\mathbf{a}}(t)$ and $\ddot{\mathbf{a}}(t)$ represent respectively the position, the rotation speed, and the acceleration of the body segments at time t, $\mathbf{T}(t)$ is the torque input for each segment, and $\mathbf{z}$ is the anthropometrical parameters reported in Table 1. The knee joint (between lower and upper leg) extension has been limited to $180°$ according to anatomy. The motion of the skeleton is computed by numerical double integration of Equation 1 to find the actual position and speed [$\mathbf{a}(t)$ and $\dot{\mathbf{a}}(t)$; cf. Figure 1].

2.2 The Dynamical Model of the Muscles

Particular care has also been taken in modeling the muscles and the neuromuscle control loop. Muscles can be considered a low-pass filter as the increase and decrease in the contraction force require a finite amount of time (Ghez, 1991). This is mainly due to the mechanical properties of cross-bridges in the muscles and can be well captured by the biomechanical models (Zangemeister, Lehman, & Stark, 1981; Kashima & Isurugi,

Table 1   The anthropometric data used in the dynamical model of the human body. They have been derived from de Leva (1996) and they refer to a man, 1.73 cm tall, weighing 73 kg; *r* is gyration radius.

Segment	Length	Mass	CM position	*r* sagittal	Inertia sagittal	*r* transversal	Inertia transversal	*r* longitudinal	Inertia longitudinal
	mm	kg	(%)	mm	kg*m^2	mm	kg*m^2	mm	kg*m^2
Head	242.9	5.0662	50.02	73.60	0.03759237	76.51	0.04062895	63.40	0.02789302
Trunk	603.3	31.7258	51.38	197.88	1.70178253	184.61	1.48115062	101.96	0.45178289
Upper trunk	242.1	11.6508	50.66	122.26	0.23856417	77.47	0.09579050	112.58	0.20226855
Medial trunk	215.5	11.9209	45.02	103.87	0.17618739	82.54	0.11124443	100.85	0.16610107
Inferior trunk	145.7	8.1541	61.15	89.61	0.08968556	80.28	0.07199055	85.53	0.08170497
Arm	281.7	1.9783	57.72	80.28	0.01746758	75.78	0.01556136	44.51	0.00536855
Forearm	268.9	1.1826	45.74	74.22	0.00892308	71.26	0.00822599	32.54	0.00171501
Hand	86.2	0.4453	79.00	54.13	0.00178757	44.22	0.00119283	34.57	0.00072884
Thigh	422.2	10.3368	40.95	138.90	0.27320680	138.90	0.27320680	62.91	0.05603666
Calf	434.0	3.1609	44.59	110.67	0.05303319	108.07	0.05056687	44.70	0.00865250
Foot	258.1	1.0001	44.15	66.33	0.00602786	63.23	0.00547808	32.00	0.00140327

1998). In these models, muscles are represented as linear dynamical systems (Figure 2a), which develop a force (contraction) as a function of the associated neural input and shortening rate. It results in a first-order model whose activation function $m(.)$ can assume two different shapes depending on the input signal $n(t)$:

$$m(t) = \begin{cases} M_o + (M_{max} - M_o)(1 - e^{-(t - t_j)/\tau})n(t) - b(\dot{\mathbf{a}}(t)) \\ \qquad \text{if } n(t) > 0 \\ M_o + (M_{min} - M_o)(1 - e^{-(t - t_j)/\tau})n(t) + b(\dot{\mathbf{a}}(t)) \\ \qquad \text{if } n(t) < 0 \end{cases} \quad (2)$$

where $b()$ is the viscous resistance and t_j is the time of transition of $n(t)$ from negative to positive or vice versa. Every time there is such a transition, M_o is defined as the tension value at that instant; rapid changes in $n(t)$, are filtered out by the low-pass filter implemented in Equation 2.

The transformation of muscle tension into joint torque is itself nonlinear as it depends on the moment arm, which varies with joint orientation (Winter, 1990). However, for small angular variations, like those in maintaining vertical posture, the moment arm can be approximated as constant.

Moreover, in the following we will not consider each muscle on its own; rather, the ensemble of muscles acting on the same joint will be treated as a single entity, a unicum (Pedotti, 1977; Winter, 1990; Lacquaniti, 1997; Anderson & Pandy, 2001). As a result the overall torque, generated by all the muscles acting on the joint i, can be represented in a general form as

$$m(t) = \begin{cases} T_{o_i} + (T_{max_i} - T_{o_i})(1 - e^{-t/\tau})n_i(t) - b(\dot{\mathbf{a}}_i(t)) \\ \qquad \text{if } n_i(t) > 0 \\ T_{o_i} + (T_{min_i} - T_{o_i})(1 - e^{-t/\tau})n_i(t) + b(\dot{\mathbf{a}}_i(t)) \\ \qquad \text{if } n_i(t) < 0 \end{cases} \quad (3)$$

The time constant, τ, was chosen according to the data in the literature (Zangemeister et al., 1981): $\tau = 20$ msec. The choice of the maximum positive, T_{max}, and negative, T_{min}, torques is particularly critical: For each
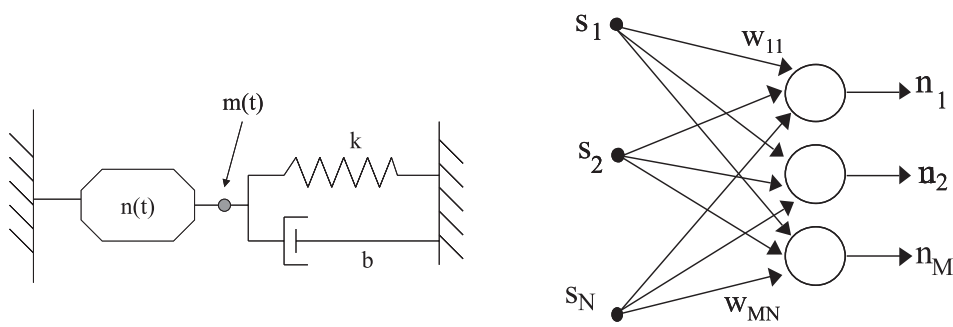
Figure 2 (a) Muscles are schematized here as linear visco-elastic models (Kashima and Isurugi, 1998). (b) The controller is implemented as a neural network. Its output depends on the value of the "synaptic weights," $\{w_{ij}^c\}$.

Table 2 The maximum extension, T_{max}, and flexion torques, T_{min}, used in this work. They were derived from Pedotti et al. (1978). CM, center of mass

	Maximum torque (extension)	Minimum torque (flexion)
Hip	384 nm	–464 nm
Knee	512 nm	–208 nm
Ankle	552 nm	–30 nm

joint these values were derived from the data of Pedotti, Krishnan, and Stark (1978), by choosing 80% of the maximum torque reported in their article. These values are plotted in Table 2.

2.3 The Postural Controller

The feedback postural controller receives as input the orientation, $\mathbf{a}(t)$, and rotation speed, $\dot{\mathbf{a}}(t)$, of the body segments (cf. Figure 1). This is a common choice in the study of motor control (e.g., Massion, 1994; Lacquaniti, 1997) and it is based on the observation that muscle and tendon receptors code stretch and stretch rate (Ghez, 1991). The vectors $\mathbf{a}(t)$ and $\dot{\mathbf{a}}(t)$ will be referred to with a single vector: the state of the system, $\mathbf{s}(t) = [\mathbf{a}(t) \mid \dot{\mathbf{a}}(t)]$. The controller transforms the state into a vector of binary signals, $\mathbf{n}(t) = [-1\ 1]$, output to the joints, and $n_i(t)$ represents a "neural spike" for the flexors or extensors muscles acting at joint i at time t. Muscles act therefore as an integrator (low-pass filter) of the "neural spike" train: To build up consistent muscle activity $\mathbf{n}(t)$ should maintain the same sign for a certain amount of time; when $\mathbf{n}(t)$ change sign often, no appreciable change in torque can be observed.

The highly nonlinear multi-input/multi-output transformation realized by the controller has been implemented here as a neural network (Figure 2b).

This is a parametric model, $h(.)$, described by the following equation:

$$\mathbf{n}(t) = h(\mathbf{a}(t),\ \dot{\mathbf{a}}(t) \mid \mathbf{w}^c) \tag{4a}$$

where $\mathbf{w}^c$ represents the parameters, called "synaptic weights": These weights tune the contribution of each input to each output variable in a nonlinear way.

To increase even more the reliability of the control function, a neural delay, Δt, has been added between the output of the controller, $\mathbf{n}(t)$, and its input, $\mathbf{s}(t)$, (Figure 1) leading to a reformulation of Equation 4a:

$$\mathbf{n}(t) = h(\mathbf{a}(t - \Delta t),\ \dot{\mathbf{a}}(t - \Delta t),\ \mathbf{w}^c) \tag{4b}$$

This delay in the controller output, $\mathbf{n}(t)$, with respect to its input, $\mathbf{s}(t)$, simulates the spinal reflex loop (Ghez, 1991; Massion, 1994): Considering that each synapse introduces a propagation delay of 1–3 ms, and that few interneurons are present in the spinal loop, a delay of 4–20 ms is usually accepted for the spinal control loop. This is in accordance with short latency reflexes recorded in reactive postural tasks (e.g., Lacquaniti, Carrozzo, & Borghese, 1993).

The neural-network controller can be viewed as an adaptive element whose behavior is specified by

the weights. These have to be properly set to maintain the upright position of the body. However a correct set of weights $\{\Delta w_{ij}{}^c\}$ is not available from the environment and it has to be *discovered*. The strategy is to increment the absolute value of the weights that make the body stand for a long time and to reduce those that lead to a fast failure. This learning procedure, called reinforcement learning (Kaelbling, Littman, & Moore, 1996; Doya, 2001), is implemented by the critic. The picture described here is biologically plausible as we suppose that human infants do not have any direct information on the desired controller behavior but do monitor when they fall down.

2.4 The Critic

The only information available to the system is when the body falls down, which can be viewed as an external reinforcement signal, *R*. This is an instantaneous signal: It could be used to correct the weights that produced the last control action but it would be of little help in finding where the control action was particularly wrong (this problem is called *temporal credit assignment*). *R* has to be transformed into continuous information suitable to modify the controller weights.

How to transform *R* into a weight update, $\{\Delta w_{ij}{}^c\}$, is the role of a second specialized unit, called the critic (Miller et al., 1990; Kaelbling et al., 1996). This unit monitors the state of the system and produces as output a scalar quantity, called internal reinforcement, *r*(*t*). To this aim, the critic first builds a risk map, *p*(*t*), of the system state: This map is not available to the system and has to be learned. In this perspective, the critic has been realized as a second neural network, which implements the following function:

$$p(t) = g(\mathbf{s}(t) \mid \mathbf{w}^r) \tag{5}$$

This risk *p*(*t*) is transformed into the internal reinforcement taking into account the risk history and the external reinforcement (cf. Equation 10). The weights of this second network, $\{w^r\}$ have to be estimated with the following criterion: They are modified in such a way that the more a state, **s**(*t*), is visited far from failure, the more secure is the control action, **n**(*t*), associated to that state. This strategy implicitly solves the credit assignment problem.

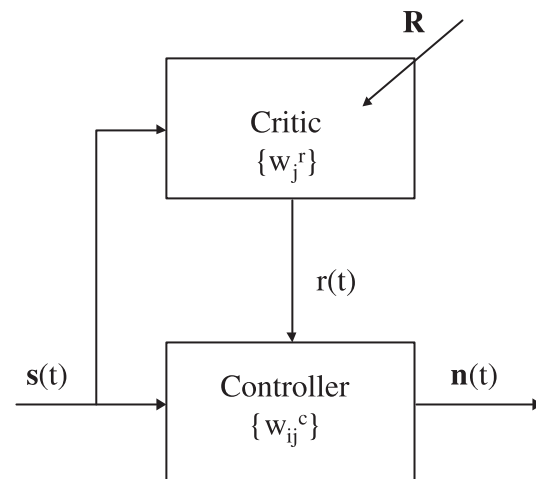The resulting control structure is hierarchical (Figure 3). The critic supervises the interaction of the

Figure 3   The hierarchical structure of the reinforcement learning controller. Both the controller and the critic are constituted of a neural network whose weights, $\{w_{ij}{}^c\}$ and $\{w_i^r\}$ have to be adequately set. To this aim, the only information available is the external reinforcement, *R*, which occurs when the controller fails to keep the body upright. From this instantaneous information, a continuos signal is estimated, the internal reinforcement *r*(*t*). This, in turns, is used to tune the controller parameters. The dimension of the controller network is N input × M output, that of the critic N input × 1 output.

controller with the environment, sending to the controller a correction signal to improve its performance. On the other side, the critic learns a high-level model of the task in the form of a risk map for the system state by monitoring the controller–environment behavior.

From the implementation point of view, the two neural networks, which realize the feedback controller (Equation 4b) and the critic (Equation 5), as well as the muscle model (Equation 3), have been implemented in Visual basic; the solution of the dynamics equation of the body motion (Equation 1) and the numerical double integration of **ä**(*t*) are carried out in the Working Model simulation environment.

3 The Implemented Reinforcement Learning Paradigm

Different machineries have been proposed to implement this hierarchical adaptive control scheme. The learning strategy considered here is a modification of the ACE/ASE (adaptive critic element/adaptive search element) model proposed by Barto, Sutton, and Anderson (1983). This remains one of the most simple and

powerful scheme and retains a higher degree of simplicity with respect to recent more elaborate schemes of reinforcement learning (Kaelbling et al., 1996; Doya, 2001). In this model the state space is subdivided into intervals creating a single vector of N state components, S, called boxes. Each box is characterized by an interval of speeds and orientations for each of the body segments. The neural networks are perceptrons whose input–output transfer function can be represented, for the controller and the critic, as

$$n_j(t) = f\left(\sum_i w_{ij}^c(t)s_i(t) + v(t)\right) \quad (6a)$$

$$p(t) = f\left(\sum_i w_i^r(t)s_i(t) + v(t)\right) \quad (6b)$$

where $f(.)$ is considered here as a threshold function: $f(x) = +1$ if $x \geq 0$ and $f(x) = -1$ if $x < 0$; $s_i(t) = 1$ only if the system state is in the ith box, $s_i(t) = 0$ otherwise. $v(t)$ is a random quantity that serves to promote *exploration* of the state space. As learning proceeds, the absolute value of the weights increases, from 0.5 to 40–50, with peaks of one order of magnitude larger for the boxes, which represent central values for the state variables. Therefore, as learning progresses, the impact of $v(t)$ on Equation 6 decreases and the system increases the *exploitation* of the most successful control actions.

The determination of the weights is done through a learning procedure as follows. At each time step t, the critic measures the actual state, $s(t)$, and outputs the corresponding internal reinforcement value $r(t)$. This is used by the controller to compute the $\{\Delta w_{ij}^c\}$ and by the critic to compute the $\{\Delta w_i^r\}$. The strategy used to update the weights is described below.

The controller weights are updated at time t according to

$$\Delta w_{ij}^c = \alpha e_{ij}(t)r(t) \quad (7)$$

where α represents the updating rate, $r(t)$ the internal reinforcement produced by the critic, and $e_{ij}^c(t)$ an eligibility measure for the state $s(t)$. This expresses the run-time estimate of the correlation between the ith state and the jth torque:

$$e_{ij}^c(t+1) = \delta e_{ij}^c(t) + (1-\delta)T_j(t)s_i(t) \quad (8)$$

where $0 \leq \delta \leq +1$. When the torque at the jth joint, T_j, produced at the ith state, s_i, is variable over time $[T_j(t)$ changes sign often in the previous time instants], the eligibility of that state is low and little updating is given to the associated weight, w_{ij}^c. The same is true if the state was seldom visited in the recent past $[s_i(.) = 0]$.

The critic internally represents the risk map of the states and uses this to compute the internal reinforcement, $r(t)$, to be delivered to the controller. The risk level, $p(t)$, of the state $s(t)$, at time t, is computed as

$$p(t) = f\left(\sum_i w_i^r(t)s_i(t)\right) \text{ with } -1 \leq p(.) \leq +1 \quad (9)$$

Then the internal reinforcement, $r(t)$, is computed from $p(t)$ as

$$r(t) = R(t) + \gamma p(t) - p(t-1) \quad (10)$$

Upon failure ($R = -1$), there is no state associated to failure, $p(t) = 0$, and the internal reinforcement, $r(t)$, will be negative and equal to $r(t) = -1 + p(t-1)$. Instead, as long as the controller succeeds in maintaining the upright posture ($R = 0$), $r(t)$ is positive when the system moves from a risky state to a less risky state $p(t) < p(t-1)$; $r(t)$ is negative vice versa. In Equation 10 a margin $\gamma(\gamma \leq 1)$ is introduced and a state is considered less risky if $\gamma p(t) < p(t-1)$.

It remains to be defined how the critic learns the risk map. This is done by again using the internal reinforcement:

$$\Delta w_i^r = \beta e_i^r(t)r(t) \quad (11)$$

where β represents the updating rate ($\beta \leq 1$) and $e_i^r(t)$ is the eligibility trace for state s_i; it is computed as (cf. Equation 8)

$$e_i^r(t+1) = \lambda e_i^r(t) + (1-\lambda)s_i(t) \quad (12)$$

where $0 \leq \lambda \leq +1$; $e_i^r(.)$ represents the frequency with which the state s_i has been visited in the recent past and it allows us to attribute a larger weight to those boxes that are occupied more recently.

4 Simulations

The ACE/ASE control system was used to control posture in the loop schematized in Figure 1, wereas

the estimate of the critic and the controller weights was carried out through the reinforcement learning strategy described in Section 3. The state of the body was discretized into 3,402 boxes with the following partition of the state variables: [–24, –12, –4, 0, +4, +12, +24] degrees for the orientation and [–∞, –50, +50, +∞] degrees/s for the rotation speed of each body segment. The body was considered out of balance, somewhat arbitrarily, when the orientation of one of the body segments fell outside the orientation range: [–24 +24] degrees. The learning parameters were set as follows: $\alpha = 1{,}000$, $\beta = 0.5$, $\gamma = 0.95$, $\delta = 0.9$, and $\lambda = 0.9$, with the following rationale. The parameters α and β have a role similar to synaptic plasticity. The large value of α determines a large modification of the controller weights (Equation 6a), such that, when the control action is successful for a certain amount of time, the $\{w_{ij}^{c}\}$ quickly reach a value that makes noise contribution to Equation 6a negligible. Lower values of α make learning slower with no improvement in the performance. β (Equation 11), instead, is more conservative: For a given state, s_i, many concordant critic outputs are required for the weights associated to that state, w_i^{r}, assuming a large value (cf. Equations 9 and 10). δ and λ represent the time constant in the decay of the eligibility trace, which is slow. The integration step was set to 4 ms, and in all the simulations, weights were updated (Equations 7 and 10) every 4 ms.

5   Results

Several experiments have been carried out and the results have been evaluated qualitatively and quantitatively in detail for eight of them. Each experiment comprises a set of trials, where each trial ends when the body falls down or when the system succeeds in maintaining upright posture. In each trial, the subject starts vertically: Although in this position the joints are aligned vertically and the body is in equilibrium, small errors introduced in the numerical solution of the dynamics (Equation 1) are sufficient to move the body out of equilibrium position; in fact, accelerations in the order of 10^{-12} degrees/s^2 are observed when the segments are in vertical position. In the first trial of each experiment the weights of the controller and of the critic are initialized randomly, and in the subsequent trials they are set to the value that thy had at the end of the previous trial. Similar results were obtained

when the weights were set to zero in the first trial of each experiment.

Learning was considered successful, somewhat arbitrarily, when the system was able to maintain upright position for at least 20 s in five consecutive trials. This was achieved, in all experiments but one, in between 700 and 1,200 trials, corresponding to about 10 hr of computational time on a Pentium III, 800 Mhz. The typical learning curve is reported in Figure 4a: As can be seen, after few trials of immediate failure, the system learns to balance for longer and longer times, until the learning curve becomes steeper. It has to be remarked that both the minimum and maximum trial duration increase with the number of trials; the variability in the intertrial duration can be ascribed to the noise in Equation 1, which promotes the exploitation of new control strategies. The learning curve shows that the learned strategy is indeed robust with respect to noise in the controller. Furthermore, analyzing the vector of the states, it can be observed that around 50% of the total states have been visited after 500 trials; at 1,000 trials most of the states have been visited at least once, while only less than 20% have never been visited; these correspond to the most extreme situations in orientation and rotation speed.

Upon learning completion, noise is removed from the controller (Equation 6a) and the updating of the controller weights through the critic output is halted. In this situation the system was able to maintain upright posture for a long time consistently (60 s of simulation were observed with no falling down).

After the controller learned to maintain the upright posture, the kinematics of the model was analyzed. The typical time course of the state variables (HAT, upper leg and lower leg orientation) is reported in Figures 5 for two different experiments.

The covariation between upper and lower leg time course is evident. The statistical analysis of the time course of the two variables gives a correlation coefficient, $r = 1$ up to the fourth decimal digit. However, a closer analysis of the time course collected in the various experiments has shown that, at some points in time, this correlation is disrupted. This can be highlighted in the homogeneous plots of Figure 6. A further insight on the kinematics time course can be gained from the relative angles (Figure 7): the ankle angle (between the foot and the lower leg), the knee angle (between the lower and the upper leg), and the hip angle (between the upper leg and the HAT). The
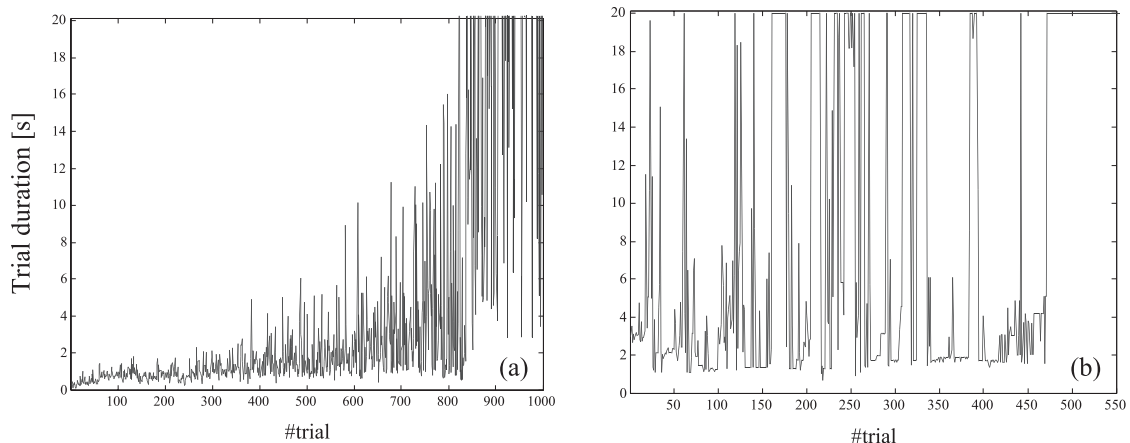
Figure 4 (a) A typical learning curve. Learning can be considered completed here after 991 trials. (b) The learning curve of a system that has to tune its control action, to balance the body when the knee was blocked. As can be seen, a few successful trials alternate with failures, until trial 476, after which the system is balanced consistently over 20 s.
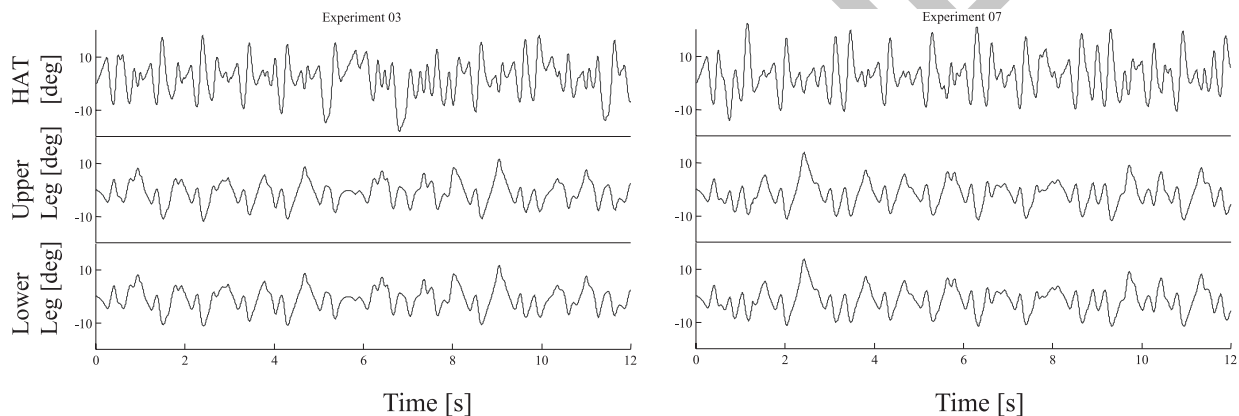

Figure 5 The time course of the HAT, the upper leg orientation, and the lower leg orientation [$\mathbf{a}(t)$ in Equation 1] plotted for two different experiments.

ankle and hip angles have a complex time course, whereas the knee angle shows a peculiar pattern: It is always at full extension (180°), except in a few instants, when peaks of a few degrees are observed. Knee flexion/extension was not present in all the experiments: In one it was absent and in two others knee flexion remained below 1°.

Furthermore, the combined rotation of the different leg segments produces an oscillatory behavior (Figures 5–7) with continuous backward/forward oscillations of the trunk. These can be visualized by plotting the time course of the center of mass in the forward/backward direction (Figure 8a) or projecting it over the sagittal plane (Figure 8b). Notice the correspondence between the peaks in knee angle (Figure 7) and in the position of the center of mass (Figure 8a), for instance at time $t \approx 5.2$ s and $t \approx 11.9$ s. The oscilla-

tions show a limited range: The maximum peak-to-peak found in the different experiments was 30 cm. Their frequency content can be clustered in two patterns: a monotonically decreasing spectrum (Figure 8c) and a spectrum with one peak (Figure 8d), located between 1.5 Hz and 2.7 Hz, depending on the experiment. In both cases, the spectrum indicates a complex behavior that cannot be accounted for by simple oscillatory models.

It should be remarked that the difference in the kinematics of the knee, the ankle and the hip, as well as the fact that the knee remains close to 180° most of the time, are the result of the learning process and are not specified a priori. Furthermore, during learning, the time course of the knee is quite variable: In particular, knee flexion/extension occurs more frequently and with a larger amplitude, up to 20°, especialy in the
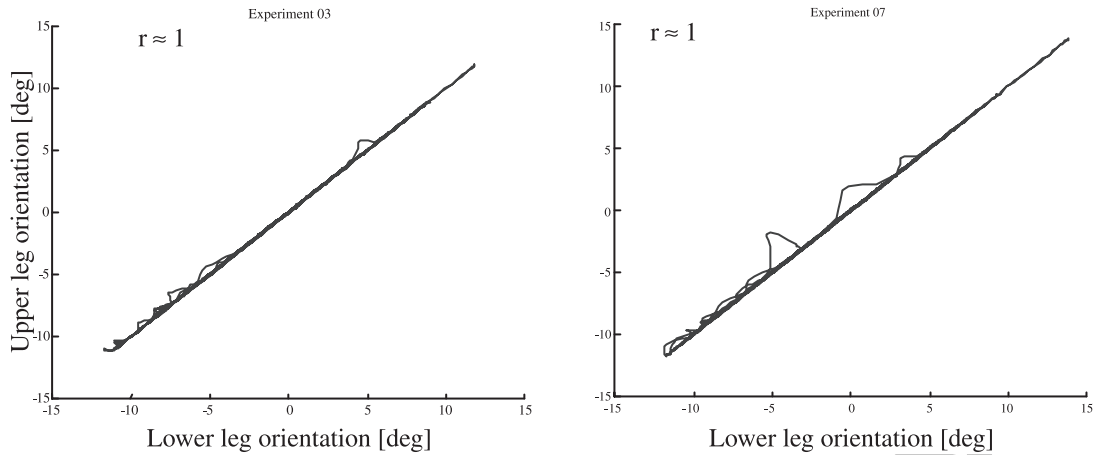
Figure 6 The upper versus lower leg orientation plotted for two different experiments.
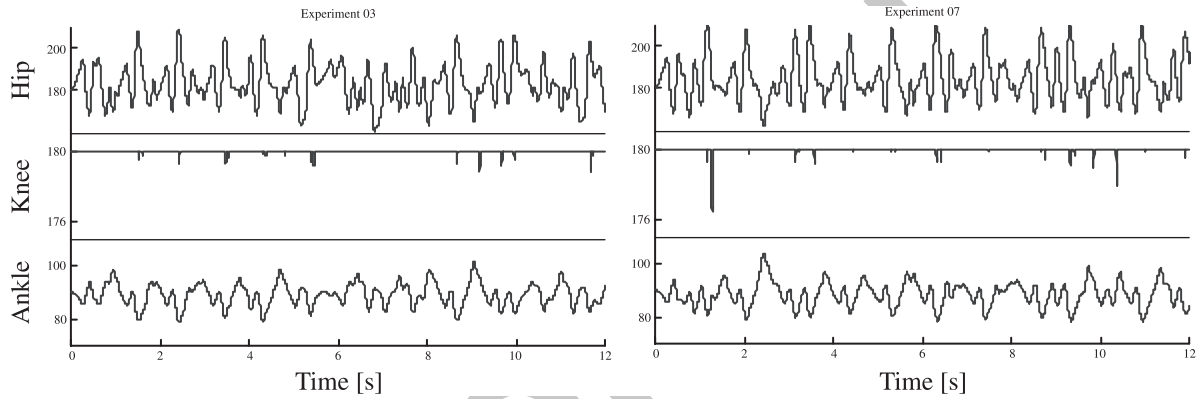

Figure 7 The time course of the anatomical angles plotted for two different experiments. Notice that the knee is completely extended most of the time. Flexion of a few degrees can be seen in certain selective instants.

first learning trials. However, a complete analysis of how these synergies emerge through learning goes beyond the scope of the present article.

5.1 Robustness and Adaptability

To test the robustness of the emerging kinematics pattern, a statistical evaluation has been carried out through the principal components analysis or Karhunen–Loeve transform (Mah et al., 1994; Borghese, Bianchi, & Lacquaniti, 1996), which is aimed in quantifying the variability of a set of signals, $\mathbf{S}(t)$. This identifies three orthogonal basis signals, which explain progressively smaller portions of the signals total variability. Mathematically the principal components are obtained through the singular value decomposition of the dispersion matrix, $\mathbf{S}(t) - \text{mean}(\mathbf{S}(t))$. In this context, principal components analysis has been applied to the

anatomical angles' time course: the hip, $h(t)$, the knee, $k(t)$ and the ankle, $a(t)$:

$$[\mathbf{U}\ \mathbf{W}\ \mathbf{V}] = \text{svd}(\mathbf{A}(t))$$

$$\text{where }\ \mathbf{A}(t) = \begin{bmatrix} h(t) - \text{mean}(h(t)) \\ k(t) - \text{mean}(k(t)) \\ a(t) - \text{mean}(a(t)) \end{bmatrix} \quad (13)$$

$\mathbf{V}$ is and orthonormal matrix, whose columns, $\mathbf{V}^j$, indicate the orientation of the principal directions; and $\mathbf{W}$ is a diagonal matrix, whose elements describe the variability in each principal component. SVD is singular valve decomposition. The jth principal component (PC) can be obtained by projecting the orientation angles over the direction $\mathbf{V}^j$ as

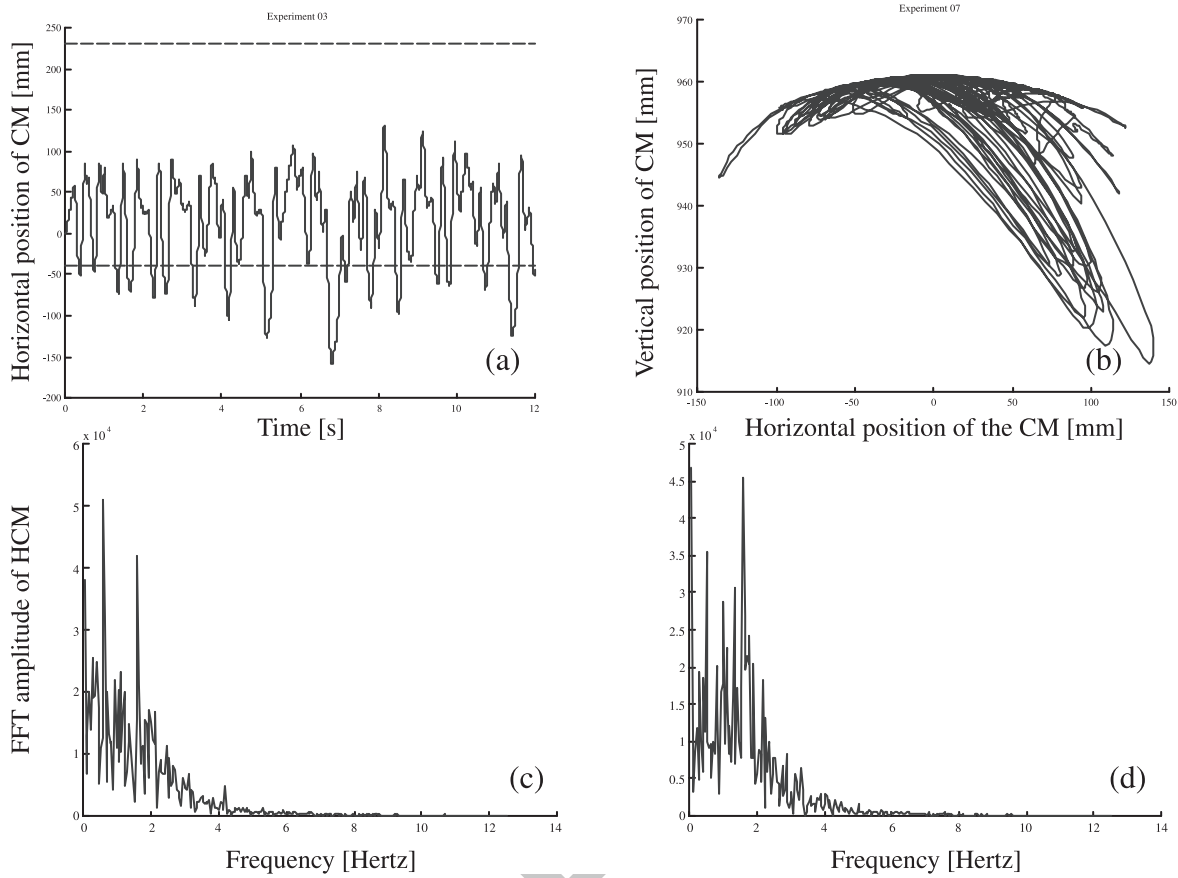$$\text{PC}^j(t) = \mathbf{V}^j * \mathbf{A}(t) \quad (14a)$$

Figure 8 (a) The time course of the forward/backward displacement of the center of mass (CM): Dashed lines indicate the feet boundary. (b) Its projection over the time course of the sagittal plane. (c, d) The amplitude of the Fast Fourier Transform (FFT) of the horizontal position of the center of mass (HCM) for two different experiments.

and, similarly, $\mathbf{A}(t)$ can be obtained from the principal components as

$$A^j(t) = V^{Tj} * \mathbf{PC}(t) \qquad (14b)$$

The direction of the three principal components and the variability explained by each of them are reported in Table 3; their time course is plotted in Figure 9 for two experiments. The first component, which alone explains, on average, 75.7% of the variability, has a time course very similar to that of the hip: It can reconstruct the hip time course with a residual error measured as mean(abs($PC1(t) - h(t)$)), of 0.655°, on average. The ankle reconstruction has a significant residual (1.524° on average, cf. Figure 9), and the second component is required to obtain an almost complete reconstruction (98.2% of the variability explained). Knee flexion/extension, instead, is completely absent and is captured by the third principal component alone: The first two components are related to the hip and ankle motion, whereas the third is related to the knee time course. This result is very robust as can be appreciated from the extremely low standard deviation reported in Table 3 for all the analyzed experiments: The standard deviation of the variations in angle orientation was 1.59° with a maximum difference between the directions of two principal components of 3.35°.

The ability of the controller to cope with sudden perturbations was tested by delivering unforeseen torque inputs to the body model. Torques of different duration (40 ms, 100 ms, 200 ms) and amplitude (10%, 30%, and 50% of the maximum torque deliverable to the joints by the ensemble of muscles) were used. The results are reported in Figure 10 and show that the controller is indeed robust against perturbations. Torques of 60% of maximal value have to be delivered to consistently destabilize the system. Torques up to 30% of the maximal value have to be delivered for

Table 3   The parameters of the singular value decomposition applied to all the analyzed experiments. PC1, PC2, and PC3 are the three principal components, where each of them is identified by its director cosines in the anatomical angles space. The "Residual" expresses the mean value of the reconstruction error when using only one principal component; the percentage of total variability taken into account by each principal component is reported in the column "Variability". Notice that the third principal component, PC3, is always oriented along the knee angle.

Experiment	PC1	PC2	PC3	Residual	Variability (%)
	0.465	0.885	0.0	0.697	72.6
1	0.0	0.0	–1.0	0.0	27.3
	–0.885	–0.465	0.0	1.325	0.1
	0.371	–0.929	0.0	0.535	75.17
2	0.0	0.0	–1.0	0.0	24.8
	–0.929	–0.371	0.0	1.341	0.03
	0.363	–0.933	0.0	0.771	77.2
3	0.009	0.0	–1.0	0.126	20.14
	–0.885	–0.371	0.0	2.003	2.66
	0.372	–0.928	0.0	0.535	75.1
4	0.0	0.0	–1.0	0.0	24.83
	–0.928	–0.372	0.0	1.340	0.07
	0.49	–0.872	0.0	0.585	77.0
5	0.0	0.0	–1.0	0.0	22.99
	–0.872	–0.49	–0.0	1.043	0.01
	0.381	–0.924	0.0	0.723	75.1
6	0.01	0.015	–1.0	0.185	20.2
	–0.924	–0.381	–0.0	1.755	4.7
	0.372	–0.928	0.0	0.612	77.7
7	0.006	0.003	–1.0	0.02	17.24
	–0.928	–0.372	–0.0	1.351	0.06
	0.355	0.355	0.0	0.748	75.5
8	0.01	0.01	–1.0	0.182	24.3
	–0.935	–0.935	–0.0	1.862	0.2
	0.402	–0.914	0.002	0.655	75.7
	(0.052)	(0.025)	(0.004)	(0.1)	(1.75)
Mean (SD)	0.004	0.003	–1.0	0.071	22.5
	(0.005)	(0.005)	(0.0)	(0.089)	(3.5)
	–0.914	–0.401	0.005	1.524	1.8
	(0.025)	(0.053)	(0.006)	(0.351)	(2.1)

400 ms to destabilize the system. Perturbations of shorter duration make the body sway (see the local peak in the HAT orientation between 1 s and 1.5 s in the central panel), but they can be absorbed by the controller. No appreciable variation can be seen when the torque perturbation is as low as 10% of maximum torques.

The adaptability of the controller to different body structures has been investigated by blocking the knee joint when learning has been completed. As can be seen by comparing Figure 4b with Figure 4a, the controller learns to balance this modified model, but the learning curve is completely different: A few consecu-
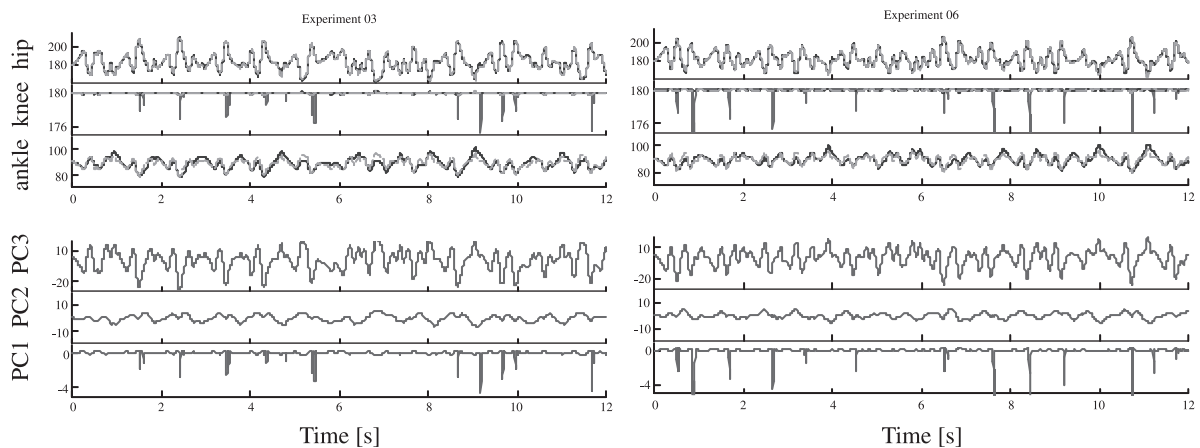
Figure 9 Principal components analysis. The reconstruction of the time course of the anatomical angles, obtained with only one principal component (PC), is plotted with a dotted line, superimposed on the true time course in the upper panels. Notice that hip reconstruction is almost complete, ankle reconstruction has a significant residual, and knee flexion/extension is completely missing. In the lower panels, the time course of the three principal components is reported.
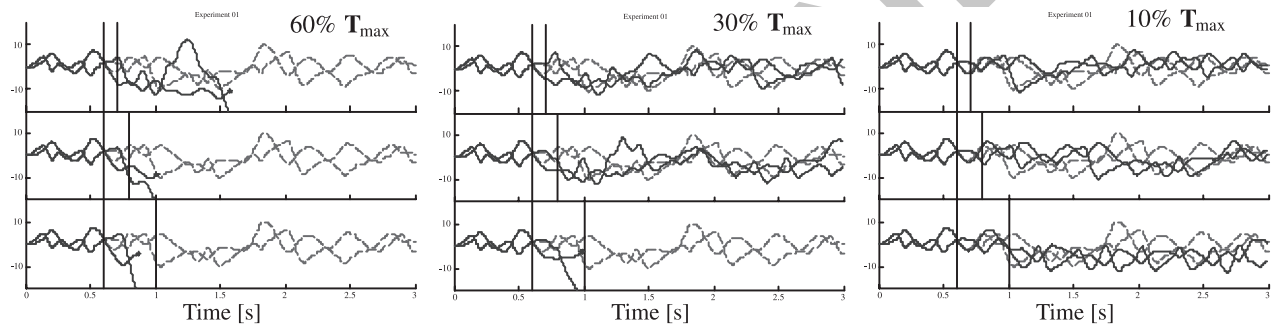

Figure 10 Effect of the perturbations. The kinematics of the HAT and of the lower leg elevation is plotted, with dotted lines, when a torque input is activated on the body model. The torque is applied at $t = 0.6$ s and lasts, from top to bottom, 100 ms, 200 ms, and 400 ms, respectively. Different level of torques have been used; from left to right: 60%, 30%, and 10% of maximum joint torques. For the sake of comparison the unperturbed kinematics from the two variables in the same experiment is plotted as a continuous line.

tive successful trials are alternated with failures until balancing becomes stable. In this experiment this was achieved at the 476th trial, but in other experiments this number was even lower. This behavior is consistent and can be studied in the context of interfering motor patterns (e.g., Shadmehr & Holcomb, 1999).

6 Discussion

Simulations show that the control system is consistently able to learn to maintain upright posture. The kinematics strategies learned in different experiments are quite similar: They all share the characteristic that the controller tends to block the knee and to act at the ankle and the hip level. This mechanism aligns the

lower and upper leg, and it can be interpreted as a control simplification since the entire leg can be controlled as a single inverse pendulum instead of a double one, made of the upper and lower leg. Therefore two kinematics synergies can be postulated. The first synergy is aimed at maintaining the upright posture by balancing the trunk over the leg, considered as a single segment rotating over the ankle: Forward/backward bending of the trunk is compensated by counterrotation of the ankle and the hip. The fact that the knee stays at the maximal extension simplifies the overall control by reducing the number of degrees of freedom to be controlled (cf. Mah et al., 1994; Lacquaniti, 1997). In the second kinematics synergy, the knee rapidly flexes and it is extended again shortly after; this movement is not correlated with hip or ankle motion as

shown by the fact that the third principal component is aligned with the knee angle (Table 3 and Figure 9), and it can be considered as a different kinematics synergy. This is activated mainly when large trunk displacement occurs, destabilizing the whole body, as can be appreciated comparing Figures 7 and 9 (cf. also Forssberg, 1999).

The overall data suggest that the hip and the ankle angles (with the knee extended) are the variables controlled in normal postural control, while knee flexion/ extension is recruited in "emergency" situations. The role of hip and ankle was first put forward by Nashner (1977), who defined a "hip strategy" and an "ankle strategy" for postural control. The two strategies were named after the joints, which were mainly in charge of compensating for perturbations, coming as tilt or displacement of the support base. In the present experiments, the controller tends to privilege a "hip strategy" as can be seen by analyzing Table 3. In fact, the first principal component, which explains most of the variability (75.7% on average), allows the reconstruction of the hip rotation in details (an average residual of 0.655° was observed), while the second principal component is required for reconstructing the ankle motion in detail (cf. also Figure 9). Instead, the ankle strategy is privileged in experiments where subjects have to learn balancing on a tilting platform, since the ankle motion become prominent in the control.

The presence of two synergies, which cope with two different situations, has also been described experimentally in other motor tasks, for instance, in gait (Borghese et al., 1996). There, a covariation between segments has been described and it has been shown that it is robust in all the gait phases and for different velocities, except in the push-off. In this gait phase the ankle angle shows a large variability, which is paralleled by the variability in the push-off force; shortly afterward the kinematics pattern resumes its stereotype.

This framework suggests an interpretation of postural control in terms of prime movers (Bouisset & Zattara, 1987). These are the elements (muscles or joints) that are active in starting a given movement, whereas the other muscles are functional in maintaining the body configuration. In this view, in normal functioning, hip and ankle joints are the prime movers in the control of posture, whereas when a large displacement of the skeleton is required, knee joint enters into play.

6.1 Methodological Issues

Two types of controllers of upright posture are discussed in the literature. According to one view, balancing could be maintained by setting the stiffness of each limb (Winter et al., 1998): By making more stiff antagonistic muscles act on a joint, the overall stiffness of that joint increases (Lacquaniti et al., 1993). Muscle stiffness can be controlled by directly regulating the stiffness (Bizzi, 1980) or the resting position (Feldman & Levin, 1995) of each muscle. In this framework, the observed kinematics is the result of the interaction of the body dynamics with the muscle dynamics. Although this pure passive control model has the merit of not being subjected to the control loop delays, it does not fit the stiffness data reported in the literature (Morasso & Schieppati, 1999), and the active control paradigm is considered in this work (e.g., Lacquaniti, 1997; Morasso & Schieppati, 1999). In this scheme, the controller outputs a set of proper spikes or torques for each joint, as a function of the whole body state, $\mathbf{s}(t)$ (Section 2.3). However, besides active control, a possible role of stiffness cannot be ruled out, and more complex controllers are required to investigate the possible interplay between stiffness and postural control.

Particular care has been put in dimensioning the torques: The data were derived from the work of Pedotti et al. (1978), where the maximum deliverable torque as a function of maximum contraction force, cross-sectional area, moment arm, and muscle lengthening were carefully estimated.

The model of the body introduced here responds to a minimalist principium: The head, the arms and the trunk are lumped into a single rigid body called the HAT in the biomechanics community (Winter, 1990). This model is used to simplify the analysis of those body movements that do not require the motion of the arms or of the head to accomplish the motor tasks (e.g., Anderson & Pandy, 2001): An increase in computational time of two orders of magnitude are observed when all the segments described in Table 1 were allowed to move. The HAT approximation has been supported by evidence, at least in gait, that kinematics data of the trunk and of the legs are not altered when arms are blocked (Borghese et al., 1996; cf. Forssberg, 1999). The release of this constraint and the analysis of a possible role of other segments, and in

particular of the arms, in learning successful postural control strategies is the goal of a future study.

6.2 Oscillatory Behavior

Oscillatory behavior has been described in the control of upright posture in adults (Collins & De Luca, 1995) and children (Forssberg, 1999). However, the oscillations' amplitude found in these experiments, although within a limited range, is larger than that observed in the real case.

This suggests that a more refined model for computing the internal reinforcement (Equation 9) has to be put forward. One possibility is to introduce a measure of energy consumption as an additional input to the critic, besides the external reinforcement, *R*. In control theory, for infinite horizon problems such as upright balancing, a measure of energy consumption widely used is some L^{loc} norm of the torques developed through time (Bryson & Ho, 1975).[2] A second possibility is that stabilization of the optical flow on the retina plays a major role in learning upright posture. It has been demonstrated that humans stabilize head and gaze during different motor tasks, possibly to get a stable representation of the external world (Pozzo, Berthoz, & Lefort, 1990). This stabilization cannot be achieved when, like in the present case, the trunk continuously moves over the support. A retinal slip error, again in some L^{loc} norm, can therefore be hypothesized as additional input to the critic. Testing these hypotheses goes beyond the scope of this article, whose main focus is on the analysis of kinematics synergies learned in a pure reinforcement learning setting.

6.3 Neurophysiology Remarks

The active controller used here is ideally located at the level of the spinal cord, which introduces a maximum delay of the order of 20 ms between sensory input and motor output (muscle contraction; Ghez, 1991). The delay is a critical element in any control system (Astrom & Witthenmark, 1989); it has been shown by simulations that a delay of the order of 50 ms is sufficient to destabilize a postural controller (Morasso & Schieppati, 1999). This same result was found here, where the controller had extreme difficulty in learning upright posture for delays larger than 20 ms. Therefore higher centers, which have response latencies of the order of 50–80 ms (brain stem or cerebellum) or >80 ms (cere-

bral cortex) cannot host the feedback controller proposed here.

However, this picture does not rule out an involvement of the higher centers in the control of posture; in particular, when perturbations can be foreseen, feedforward control schemes based on internal models can be used (Wolpert & Kawato, 1998). Moreover, higher centers, and in particular the basal ganglia, can host the critic module, and the basal ganglia–cortex loop can host the reinforcement learning machinery (Doya, 2001). In fact, the critic does not need to operate synchronously with the controller: It is only required that its output, the internal reinforcement, is delivered to the right state/control pair, although this signal can be delayed in time.

7   Conclusion and Future Work

A new methodology to study learning complex human behavior has been presented here. Overall, the results show that the anatomical arrangement of the skeleton is sufficient to shape a postural control, robust against torque perturbations and noise, and flexible enough to adapt to changes in the body model in a short time. Moreover, the learned kinematics closely resembles the data reported in the literature; it emerges from the interaction with the environment, through trial-and-error; no a priori information is given as a hard-wired property of the control system.

The system can be extended in several ways by modifying the body model or the controller. The HAT hypothesis can be released and all is macro-segments in Table 1, and in particular the arms, can be allowed to move; this might highlight finer control mechanisms in balancing and allows studying the influence of the different body segments on learning upright posture. As far as the controller is concerned, it can be made more complex, by using multiple neural networks in parallel, each with a different input/output function, competing to produce the correct control action sequence (cf. Jordan & Jacobs, 1994). In an interesting experiment, one neural network can be the one used in this article, and a second one can implement a stiffness control, receiving as input the position and rotation velocity of the limbs and outputting for a given joint a signal that activates opposing pairs of muscles acting on that joint. This kind of controller might help in elucidating the relative role of passive

stiffness controller versus active torque controller in maintaining upright posture, which is still under debate (Winter et al., 1998; Morasso & Schieppati, 1999).

The methodology introduced here can be of interest in fields other than motor control. In robotics it can be useful to teach complex robots and in the new field of digital animation it can be fruitfully adopted to create digital stunts, which would replace the costly motion capture sessions made with real stunts (e.g., Popovic, Seitz, Erdmann, Popovic, & Witkin, 2000).

Notes

1 It should be remarked that maintaining the center of mass inside the support base is the definition itself of equilibrium and not an explanation!

2 Given a sequence $X = \{x_k \in R^p\}$ with $-\infty < k < +\infty$, its L^{loc} norm is defined over a subset $Y \subseteq X$, as $\sqrt[p]{\sum_{k:x_k \in Y} |x_k|^p}$.

Acknowledgements

References

Anderson, F. C., & Pandy, M. G. (2001). Static and dynamic optimization solutions for gait are practically equivalent. *Journal of Biomechanics, 34*, 153–161.

Astrom, K. J., & Wittenmark, B. (1989). *Adaptive control*. New York: Addison-Wesley.

Barto, A. G., Sutton, R. S., & Anderson, C. W. (1983). Neuron-like adaptive elements that can solve difficult learning problems. *IEEE Transactions on Systems Man and Cybernetics, 13*, 834–846.

Bernstein, N. (1967). *The co-ordination and regulation of movement*. Oxford: Pergamon Press.

Bizzi, E. (1980). Central and peripheral mechanisms in motor control. In G. E. Stelmach & J. Requen (Eds.), *Tutorials in motor behavior* (pp. 131–144). Amsterdam: North Holland.

Borghese, N. A., Bianchi, L., & Lacquaniti, F. (1996). Kinematic determinants of human locomotion. *Journal of Physiology, 494*, 863–879.

Bouisset, S., & Zattara, M. (1987). Biomechanical study of the programming of anticipatory postural adjustments associated with voluntary movement. *Journal of Biomechanics, 20*, 735–742.

Bryson, A. E., & Ho, Y. C. (1975). *Applied optimal control*. Wahsington, DC: Hemisphere.

Collins, J. J., & De Luca, C. J. (1994). Upright, correlated random walks: A statistical-biomechanics approach to the human postural control system. *Chaos, 5*(1), pp. 57–63.

De Leva, P. (1996). Joint center longitudinal positions computed from a selected subset of Chandler's data. *Journal of Biomechanics 29*(9), 1231–1233.

Doya, K. (2001). Reinforcement learning in continuous time and space. *Neural Computation*, *12*, 219–245.

Feldman, A. G., & Levin, M. F. (1995). The origin and use of positional frames of reference in motor control. *Behavioral and Brain Sciences, 18*, 723–806.

Forssberg, H. (1999). Neural control of human motor development. *Current Opinion in Neurobiology, 9*, 676–682.

Ghez, C. (1991). Posture. In E. Kandel, J. Schwartz, & T. Jessell (Eds.), *Principles of neural science* (3rd edn., part VI, pp. 534–659). Amsterdam: Elsevier.

Gurfinkel, V. S., Levik, Y. S., Popov, K. E., Smetanin, B. M., & Shilkov, V. Y. (1988). Body scheme in the control of postural activity. In V. S. Gurfinkel, M. E. Ioffe, J. Massion, & J. P. Rolle (Eds.), *Stance and motion: Facts and concepts* (pp. 147–185). New York: Plenum Press.

Jordan, M., & Jacobs, R. A. (1994). Hierarchical mixtures of experts and the EM algorithm. *Neural Computation, 6*, 181–241.

Kaelbling, L. P., Littman, M. L., & Moore, A. W. (1996). Reinforcement learning: A survey. *Journal of Artificial Intelligence Research*, *4*, 237–285.

Kashima, T., & Isurugi, Y. (1998). Trajectory formation based on physiological characteristics of skeletal muscles. *Biological Cybernetics*, *78*, 413–422.

Lacquaniti, F. (1997). Frames of reference in sensorimotor coordination. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology* (Vol. 11, pp. 27–64). Amsterdam: Elsevier.

Lacquaniti, F., Carrozzo, M., & Borghese, N. A. (1993). Time-varying mechanical behavior of multi-jointed arm in man. *Journal of Neurohysiology, 69*(5), 1443–1464.

Lacquaniti, F., & Maioli, C. (1995). In M. Arbib (Ed.), *Handbook of brain theory and neural networks* (pp. 567–568). Cambridge, MA: MIT Press.

Mah, C. D., Hulliger M., Lee, R. G. & O'Callaghan, I. S. (1994). Quantitative analysis of human movement synergies: Constructive pattern analysis for gait. *Journal of Motor Behavior*, *26*, 83–102.

Massion, J. (1994). Postural contol system. *Current Opinion in Neurobiology, 4*, 877–887.

Miller, T., Sutton, R., & Werbos, P. (Eds). (1990). *Neural networks for control*. Cambridge, MA: MIT Press.

Morasso, P., & Schieppati, M. (1999). Can muscle stiffness alone stabilize upright standing? *Journal of Neurophysiology, 83*, 1622–1626.

Mounchino, L., Cincera, M., Fabre, J. C., Assainte, C., Amblard, B., Pedotti, A., & Massion, J. (1996). Is the regulation of center of mass maintained during leg movement under microgravity conditions? *Journal of Neurophyisiology, 76*, 1212–1223.

Nashner, L. M. (1977). Fixed patterns of rapid postural responses among leg muscles during stance. *Experimental Brain Resesarch, 30*, 13–24.

Pedotti, A. (1977). A study of motor coordination and neuromuscular activities in human locomotion. *Biological Cybernetics, 26*, 53–62.

Pedotti, A., Krishnan, V. V., & Stark, L. (1978). Optimization of muscle-force sequencing in human locomotion. *Mathematical Bioscience, 38*, 57–76.

Popovic, J., Seitz, S. M., Erdmann, M., Popovic, Z., & Witkin, A. P. (2000). Interactive manipulation of rigid body simulations. *Proceedings of SIGGRAPH 2000* (pp. 209–218). Washington, DC: Addison-Wesley.

Pozzo, T., Berthoz, A., & Lefort, L. (1990). Head stabilization during various locomotor tasks in humans. *Experimental Brain Research*, *82*, 97–106.

Shadmehr, R., & Holcomb, H. H. (1999). Inhibitory control of competing motor memories. *Experimental Brain Research*, *126*, 235–251.

Winter, D. (1990). *Biomechanics and motor control of human movement*. Waterloo, Canada: Wiley InterScience.

Winter, D., Patla, A., Prince, F., Ishac, M, & Gielo-Perzak, K. (1998). Stiffness control of balance in quiet standing. *Journal on Neurophysiology*, *80*, 1211–1221.

Wolpert, D., & Kawato, M. (1998). Internal models of the cerebellum. *Trends in Cognitive Science, 2*, 338–347.

Zangemeister, W. H., Lehman, S., & Stark, L. (1981). Simulation of head movement trajectories: Model and fit to main sequence. *Biological Cybernetics*, *41*, 19–32.

Zatsiorsky, V. M., Raitsin, L. M., Seluyanov, V. N., Aruin, A. S., & Prilutzky, B. J. (1993). Biomechanical characteristics of the human body. In W. Baumann (Ed.), *Biomechanics and performance in sport* (pp. 71–83). Washington, DC: Addison-Wesley.

About the Author

N. Alberto Borghese is associate professor at the Department of Computer Science of the University of Milano, where he teaches intelligent systems and digital animation and directs the laboratory of motion analysis and virtual reality. He graduated with full marks and honors from Politecnico of Milano in 1984–1985; he was visiting scholar at Center for Neural Engineering of USC in 1991, at the Department of Electrical Engineering of Caltech in 1992, and at the Department of Motion Capture of Electronic Arts, Canada in 2000. His research interests include quantitative human motion analysis, modeling and synthesis in virtual reality, and artificial learning systems, areas in which he has authored more than 30 papers in refereed journals.

Andrea Calvi is currently a masters student in cooperation and development, at the European School for Advanced Studies at the University of Pavia, Italy. He graduated from Politecnico of Milano in 2000–2001. His research interests are in both intelligent systems and techniques for developing the economics of Third World countries.