A biped static balance control and torque pattern learning under unknown periodic external forces

Satoshi Ito *, Tomohiro Kashima, Minoru Sasaki

Department of Human and Information Systems, Faculty of Engineering, Gifu University, Yanagido 1-1, Gifu 501-1193, Japan

ABSTRACT

This paper addresses a biped balancing task in which an unknown external force is exerted, using the so-called 'ankle strategy' model. When an external force is periodic, a human adaptively maintains the balance, next learns how much force should be produced at the ankle joint from its repeatability, and finally memorized it as a motion pattern. To acquire motion patterns with balancing, we propose a control and learning method: as the control method, we adopt ground reaction force feedback to cope with an uncertain external force, while, as the learning method, we introduce a motion pattern generator that memorizes the torque pattern of the ankle joint by use of Fourier series expansion. In this learning process, the period estimation of the external force is crucial; this estimation is achieved based on local autocorrelation of joint trajectories. Computer simulations and robot experiments show effective control and learning results with respect to unknown periodic external forces.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

A kind of intelligent motor behaviors of animals is observed in a learning of motion patterns by adapting to unknown environment. Even my pet dog, for instance, not only walks in my cluttered house without tripping but also changes his walking or running pattern according to the situation. Such behaviors are not easy to achieve as robot behaviors, because, in conventional robot controls, the motion patterns are programmed in advance by assuming environmental conditions. The programmed robot behaviors are not assured in unknown or variable environments.

Two types of abilities are found in the above dog behavior. One is an ability to learn motions as a special pattern appropriate to a steady environmental condition (walking is switched to running). The other is to stabilize the posture or movement to deal with a rapid disturbance from the environment (walking in the cluttered house where foot placement may be slightly different at each step). This paper treats a scheme of motor control and learning from these two points of view.

To make the problem as simple as possible, the balancing problem as shown in Fig. 1 is considered as an example. In this situation, a human stands on a floor where the slope changes periodically at a slow speed, e.g., on a large boat in the wild sea. Normally, humans can adjust their standing position with respect to the slope of the ground by changing the sway angle, in other words, the joint angle of the ankle. This balancing method is especially called ankle strategy (Horak and Nashner, 1986), although some other strategies, mechanisms or control methods are proposed in the human standing (Winter, 1995; Alexandrov et al., 1998; Rietdyk et al., 1999; Van Ooteghem et al., 2008) or biomechanical/robotic model (Gorce and Vanel, 1997; Hof, 2007; Mergner et al., 2009). In addition, if this situation continues for a long time, humans can learn the operation of the ankle as a periodic motion pattern based on this periodicity. The object of this paper is how to achieve such a balancing behavior by artificial machines like robots rather than to elucidate a neuronal mechanism whereby humans accomplish such an adaptive behavior, although the problem originates from a consideration of a control/learning mechanism in the behaviors of biological systems.

Balance control methods are often discussed in the field of robotics, especially for walking robots. Here, the main issue is rather motion generation based on the zero moment point (ZMP) (Vukobratovic et al., 1990), a point on the ground around which the moment of inertial force and gravity are balanced. In this method, reference trajectories of the joints or the body's center of gravity (CoG) are calculated in advance as a motion pattern so that the ZMP is kept beneath the foot support. Then the trajectories are reproduced by position control during actual walking (Takanishi et al., 1989; Hirai et al., 1998; Mitobe et al., 2001). On-line generation (Kajita and Tani, 1996; Nishiwaki et al., 2002; Sugihara et al., 2002; Behnke, 2006) or on-line modification (Huang et al., 2000; Wollherr and Buss, 2004; Lee et al., 2005; Prahlad et al., 2007) of the trajectories have been proposed to adapt to changes in environmental conditions. However, the discussions of periodic pattern learning are not yet sufficiently advanced to support the development of walking robots. As for
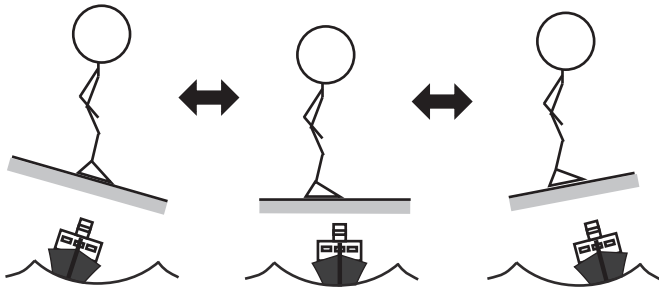
Fig. 1. Balance control affected by periodic external force.

static balance, the stability of the upright posture is analyzed (Napoleon and Sampei, 2002). The acquisition of biped dynamics is described with a neurophysiological model (Nakayama and Kimura, 2004). Reinforcement learning is applied (Borghese and Calvi, 2003), also to a stand-up behavior that requires a balancing task (Morimoto and Doya, 2001). Human static balance is measured to clarify its adaptive characteristics (Nashner, 1976; Priplata et al., 2002; Lockhart and Ting, 2007). Regarding to the periodic motion learning, a method based on the controllers containing oscillators, called by CPG, have been proposed (Nishii, 1999; Ishiguro et al., 2003; Ijspeert, 2008; Endo et al., 2008). However, control and pattern learning schemes of periodic motion based on the on-line balance have not been sufficiently discussed.

Thus, this paper deals with a problem such as that in Fig. 1 that contains both control and learning factors. We have already proposed a control and learning method for a special case where the period of an environmental alternation is given beforehand (Ito et al., 2005). In this paper, this method is extended to the case in which its period is also unknown. For this extension, estimation of the period is required. After the concept of our control scheme is explained in Section 2, a control and learning method including period estimation is formulated in Section 3. In Section 4, the effects of our scheme are confirmed by simulations, and it is applied to the motion control of actual robot in Section 5. In Section 6, concepts and assumptions in this paper are reconsidered and its advantages as well as remained problems are discussed. Finally, this paper is concluded in Section 7.

2. Control and learning scheme for static balance

2.1. Strategy

One of our future goals is to apply the balancing and learning method to the locomotion pattern learning. The walking is periodic motion and thus the balance disturbances caused by the inertial force of walking also become periodic. Thus, motion pattern learning under periodic external force, such as in Fig. 1, is compatible to the locomotion pattern learning and is applicable to the design of the desired trajectories in the locomotion pattern. Then, an irregular environment, such as seen in a cluttered house, produces non-periodic disturbances. Under the plan in this paper, such non-periodic disturbance is compensated by the feedback control based on ground reaction forces. Simultaneously, an adequate motion pattern is learned into the motion pattern generator to automatically cope with the periodic forces from walking motions. As the first step toward this goal, we consider here the static balance control and the motion pattern learning with respect to periodic external forces.

In the human balance control, ankle strategy is typically observed for small disturbance, in which balance is kept by ankle

joint operation without moving any other joints such as knee and hip (Horak and Nashner, 1986). In the case illustrated in Fig. 1, the tilting motion of the floor on a large boat is slow, i.e., small in short time. Thus, it is reasonable to assume an ankle strategy to consider a control and learning scheme in biped balance control.

As mentioned in Section 1, adaptability should enable the adjustment of motion to environmental changes on different timescales. In a balancing task, this is expressed as the response to a temporary external force as a short-term fluctuation and to an ongoing, here periodic, external force as a long-term transition. These two kinds of external force cause a contradiction in motion pattern generation: In the former case, the motion pattern should be unchanged (i.e., stabilized), because the external force is regarded as the disturbance. In the latter case, the motion patterns should sometimes be changed, i.e., restructured or switched, to be appropriate to the new environment.

To cope with this contradiction, the following two characteristics are introduced to a control and learning scheme: feedback information regarding the ground reaction force and a feedforward controller as the motion pattern generator. This scheme is illustrated in Fig. 2. Here, the following scenario is postulated: The ground reaction force is informative for balance. The center of action of all the ground reaction forces is called the center of pressure (CoP) and coincides with the ZMP (Goswami, 1999). Thus, balance control against the temporary external force is constructed based on the feedback information regarding ground reaction force, as shown in Fig. 2(a). Now, assume that the external force is exerted in a periodic manner, as in Fig. 1. Then the action of the balancing motion also becomes periodic. Based on this periodicity the motion pattern is being stored in memory during the learning process. Here, the torque trajectories are considered to be the learning factor as a motion pattern. A benefit of learning is balance maintenance without information on the ground reaction forces, which generally provide significant feedback for balancing. This process is illustrated in Fig. 2(b).

2.2. Static balance model

To achieve the scenario in the above section, a simple link model, as shown in Fig. 3(a) is considered. The ankle joint acts mainly against the small disturbance, which is called ankle strategy (Horak and Nashner, 1986). From this point of view, an inverted pendulum with small foot support is introduced with the
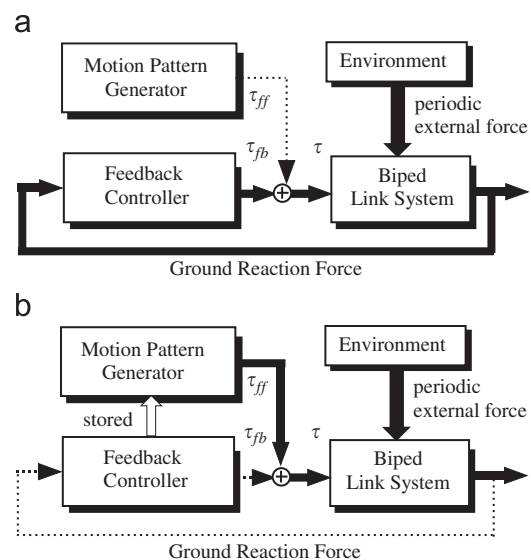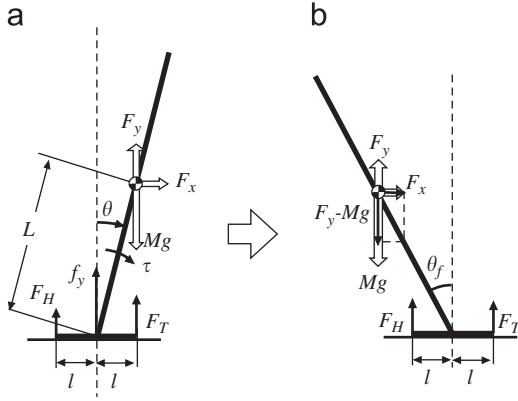

Fig. 2. Control and learning scheme. (a) Before learning. (b) After learning.

Fig. 3. Model and stationary posture of the control law (7). (a) Link model. (b) Stationary state.

following assumptions: the motion is restricted within the sagittal plane. The ankle joint can generate torque τ and its deviation θ and velocity $\dot{\theta}$ is detectable. The foot contacts ground only at two points, where the vertical component of the ground reaction force, F_T and F_H, is detectable. The ankle joint is located at the center of the foot and enough low position.

The dynamics of the inverted pendulum is given by

$$I\ddot{\theta} = MLg\sin\theta + F_x L\cos\theta - F_y L\sin\theta + \tau = F_A L\sin(\theta - \theta_f) + \tau, \qquad (1)$$

where M is its mass, I is its moment of inertia around the joint, L is the distance from the joint to the center of mass (CoM), g is the gravity acceleration, F_x and F_y are unknown external forces. F_A and θ_f are given as follows:

$$F_A = \sqrt{(Mg - F_y)^2 + F_x^2}, \qquad (2)$$

$$\tan\theta_f = -\frac{F_x}{Mg - F_y}. \qquad (3)$$

The foot support, on the other hand, never move when the balance is maintained, implying that the moment balance equations are obtained instead of the motion equation. They are given as

$$F_H = \frac{1}{2\ell}\tau + \frac{1}{2}mg + \frac{1}{2}f_y, \qquad (4)$$

$$F_T = -\frac{1}{2\ell}\tau + \frac{1}{2}mg + \frac{1}{2}f_y. \qquad (5)$$

Here, m is its mass, ℓ is the distance from the joint to the foot end, f_y is an interacting force from the inverted pendulum.

2.3. Methods

The scheme is realized using the concept of feedforward and feedback control. Here, feedback means that the system uses the signal from the ground reaction force, and feedforward means that it does not contains its sensory feedback.

In the next section, the control law is first constructed for the ankle joint torque, τ_{fb}. For periodic external forces, τ_{fb} also becomes periodic. Based on this periodicity, the torque pattern is learned; i.e., the trajectories of τ_{fb} are memorized during its repetitive generation. The motion pattern generator, acting as a feedforward controller, stores this pattern. The output torque of the feedforward controller is denoted by τ_{ff}.

The total torque τ is the sum of τ_{ff} and τ_{fb}:

$$\tau = \tau_{ff} + \tau_{fb}. \qquad (6)$$

Assume that $\tau_{ff} = 0$ initially holds. Then τ is composed entirely of the τ_{fb} component. As τ_{ff} is being learned, τ_{fb} is gradually decreasing. Finally, τ_{fb} is copied to τ_{ff}, i.e., replaced by τ_{ff}. In the next section, this process is formulated using the static balance model in Section 2.2.

3. Formulation

3.1. Adaptive balance maintenance

Not only to keep F_T and F_H positive but also to make them equal is a reasonable description of the control purpose for balance control, since the CoP or ZMP is regulated to the center of the foot. Then, the stability margin (McGhee and Frank, 1968) is maximum, implying that this posture is maintainable against any external forces. To achieve such situation, the next control law is applied:

$$\tau_{fb} = -K_d\dot{\theta} - K_p\theta + K_f \int (F_H - F_T)\, dt. \qquad (7)$$

Proposition. *Define the control law as (7), i.e., $\tau = \tau_{fb}$ for the dynamical system (1), (4) and (5). At the stationary state, $\theta = \theta_f$ as well as $F_T = F_H$ hold for constants F_x and F_y.*

Proof. At first, a new variable τ_f is defined as

$$\tau_f = \int (F_H - F_T)\, dt. \qquad (8)$$

Then, the control law (7) becomes

$$\tau = -K_d\dot{\theta} - K_p\theta + K_f\tau_f. \qquad (9)$$

In addition, differentiating (8) and then substituting (5) and (4), the following equation is obtained:

$$\dot{\tau}_f = \frac{1}{\ell}\tau. \qquad (10)$$

Now, θ, $\dot{\theta}$ and τ_f are regarded as a state variable. An equilibrium point $(\bar{\theta}, \bar{\dot{\theta}}, \bar{\tau}_f)$ of the dynamics (1) and (10) with control law (9) is obtained by setting the differential term zero. It is given as

$$(\bar{\theta}, \bar{\dot{\theta}}, \bar{\tau}_f) = \left(\theta_f, 0, \frac{K_p}{K_f}(\theta_f - \theta_d)\right). \qquad (11)$$

At this state, $\tau = 0$ holds according to (9) and (11), indicating that $F_H = F_T$. The local stability of this equilibrium point is ensured from the controllability of the linearized dynamics around this point (Ito and Kawasaki, 2005). $\square$

The posture at the stationary state is depicted in Fig. 3(b). At this posture, the moment of the external force and gravity is balanced. Namely, no torque is required at the ankle joint and thus the effective posture maintenance is achieved. Furthermore, the dependence of the stationary state θ_f on F_x and F_y indicates that the posture adaptively change with the external forces. The local stability is the same as the actual human motion: humans certainly tumble if they are pushed by a large force.

3.2. Learning of periodic pattern

This section treats the learning of periodic motion pattern as torque trajectory, when the periodic external force is exerted. Here, the period of the external force T_e is assumed to be known. The case in which its period is unknown will be discussed in the next section.

In the static balance model with ankle strategy, the ankle joint is the sole actuator. Thus, the ankle joint torque is stored to the

motion pattern generator followed by the method in Section 2.3. Namely, a control and learning law is defined as follows:

- The control law τ is defined as the summation of the feedforward torque τ_{ff} and feedback torque τ_{fb}:

$$\tau = \tau_{ff} + \tau_{fb}. \tag{12}$$

- τ_{ff} is calculated by the next equation:

$$\tau_{ff} = Y_r \hat{\phi}. \tag{13}$$

Here,

$$Y_r = [\ddot{\theta}_r, S, S_0 C, C_0 C, S_0 S, C_0 S, \ldots, S_n C, C_n C, S_n S, C_n S], \tag{14}$$

where $\ddot{\theta}_r$ is calculated from the derivative of $\dot{\theta}_r$ defined as

$$\dot{\theta}_r = -\frac{K_p}{K_d} \theta \tag{15}$$

and $S_k = \sin k\omega_e t$, $C_k = \cos k\omega_e t$, $\omega_e = 2\pi/T_e$, $C = \cos\theta$ and $S = \sin\theta$. Note that the definition of θ_r is not required in this control and learning method. $\hat{\phi}$ is, on the other hand, an estimate of the unknown parameter vector ϕ defined as

$$\phi = K_I \sigma, \tag{16}$$

where

$$K_I = \frac{K_d \ell}{K_d \ell - K_f I} \tag{17}$$

and

$$\sigma = [I, -MgL, -L\alpha_0^{(x)}, -L\beta_0^{(x)}, L\alpha_0^{(y)}, L\beta_0^{(y)}, \ldots, -L\alpha_n^{(x)}, -L\beta_n^{(x)}, L\alpha_n^{(y)}, L\beta_n^{(y)}]^T. \tag{18}$$

$\hat{\phi}$ is updated in the next learning rule.

- The learning of the feedforward torque τ_{ff} is described by the update of $\hat{\phi}$ as

$$\dot{\hat{\phi}} = -\Gamma Y_r^T s. \tag{19}$$

Here, Γ is a positive diagonal matrix.

- For the feedback torque τ_{fb}, the definition (7) is utilized:

$$\tau_{fb} = -K_d \dot{\theta} - K_p \theta + K_f \int (F_H - F_T) \, dt = -K_d s, \tag{20}$$

$$s = \dot{\theta} - \dot{\theta}_r - \frac{K_f}{K_d} \int (F_H - F_T) \, dt. \tag{21}$$

Proposition. *Assume that the period of the external force T_e is known. If a control and learning law is given as (12) and (19), τ_{ff} is constructed, in other words, $\hat{\phi}$ is learned, so that τ_{fb} goes to zero.*

Proof. Because of the periodicity, the external forces F_x and F_y are expanded to the Fourier series as follows:

$$F_x = \sum_k^n \{\alpha_k^{(x)} S_k + \beta_k^{(x)} C_k\}, \tag{22}$$

$$F_y = \sum_k^n \{\alpha_k^{(y)} S_k + \beta_k^{(y)} C_k\}. \tag{23}$$

Substituting them into (1), the next equation is obtained:

$$I\ddot{\theta} - MLgS - \sum_k^n \{\alpha_k^{(x)} S_k + \beta_k^{(x)} C_k\} LC + \sum_k^n \{\alpha_k^{(y)} S_k + \beta_k^{(y)} C_k\} LS = \tau. \tag{24}$$

Using the definition of Y_r, i.e., (14)

$$I\ddot{\theta}_r - MLgS - \sum_k^n \{\alpha_k^{(x)} S_k + \beta_k^{(x)} C_k\} LC + \sum_k^n \{\alpha_k^{(y)} S_k + \beta_k^{(y)} C_k\} LS = Y_r \sigma \tag{25}$$

is obtained. Subtract (25) from (24), we obtain

$$I(\ddot{\theta} - \ddot{\theta}_r) = \tau - Y_r \sigma. \tag{26}$$

Furthermore, from (10), we also get

$$\frac{IK_f}{K_d} \dot{\tau}_f = \frac{IK_f}{\ell K_d} \tau. \tag{27}$$

Subtract (27) from (26),

$$I\left(\ddot{\theta} - \ddot{\theta}_r - \frac{K_f}{K_d} \dot{\tau}_f\right) = \left(1 - \frac{IK_f}{\ell K_d}\right)\tau - Y_r \sigma \tag{28}$$

i.e.,

$$IK_I \dot{s} = \tau - Y_r \phi \tag{29}$$

is satisfied. Now, consider the following positive definite function:

$$V = \frac{1}{2}(K_I I s^2 + \overline{\phi}^T \Gamma^{-1} \overline{\phi}), \tag{30}$$

where

$$\overline{\phi} = \hat{\phi} - \phi. \tag{31}$$

Note that $\dot{\overline{\phi}} = \dot{\hat{\phi}}$ since ϕ is constant. Then,

$$\dot{V} = K_I I s \dot{s} + \dot{\overline{\phi}}^T \Gamma^{-1} \overline{\phi}$$
$$= (\tau - Y_r \phi)s + \dot{\hat{\phi}}^T \Gamma^{-1} \overline{\phi} = (Y_r \hat{\phi} - K_d s - Y_r \phi)s + (-\Gamma Y_r^T s)^T \Gamma^{-1} \overline{\phi}$$
$$= -K_d s^2 + Y_r \overline{\phi} s - s^T Y_r \overline{\phi} = -K_d s^2 \leq 0 \tag{32}$$

and

$$\ddot{V} = -2K_d s \dot{s}. \tag{33}$$

$V > 0$ and $\dot{V} \leq 0$ in $(s, \overline{\phi}) \neq (0, 0)$ ensure the boundedness of V, i.e., the boundedness of the s and $\overline{\phi}$. Next, the boundedness of s implies the boundedness of θ and $\dot{\theta}$, while the boundedness of $\overline{\phi}$ does the boundedness of $\hat{\phi}$. Thus, Y_r becomes bounded. Finally, $\dot{s}$ is bounded since the left-hand side of (29) is bounded. Accordingly, $\ddot{V}$ is bounded because of the boundedness of s and $\dot{s}$ in (33).

The boundedness of $\ddot{V}$ ensures that $\dot{V} \to 0$ when $t \to \infty$ based on the Lyapunov-like lemma (Slotine et al., 1991). This implies $s \to 0$. Accordingly, the update of $\hat{\phi}$ leads τ_{fb} to zero. $\quad\square$

3.3. Period estimation of external force

The formulation of the previous section requires that the period of the external force is known. However, the period is generally unknown beforehand or varies with the situation. Thus, its estimation is discussed in this section. This estimation is based on the observation of the ankle joint motion: the ankle joint moves with the same period as that of the periodic external force if the balance is maintained. Because the period is uniquely determined from the angular frequency, ω_e is estimated instead of the period T_e.

For this estimation, the effect of low pass filter was utilized in our previous paper (Ito et al., 2005). However, the generated sinusoidal function was somewhat deformed, which does not always lead to the good learning result. This is why the other method based on the local autocorrelation is introduced.

The angular frequency is estimated by the following steps.

1. The trajectory of the ankle joint angle is stored back in time during the interval T_s from t_0. This trajectory is put to $\theta_s(t, t_0)$:

$$\theta_s(t, t_0) = \theta(t_0 + t) \quad (-T_s \leq t \leq 0). \tag{34}$$

2. Using this $\theta_s(t, t_0)$, the local autocorrelation $r(t)$ is calculated by

$$r(t) = \frac{\langle \theta_s(t, t_0), \theta(t) \rangle_{T_s}}{\langle \theta(t), \theta(t) \rangle_{T_s} \cdot \langle \theta_s(t, t_0), \theta_s(t, t_0) \rangle_{T_s}}. \tag{35}$$
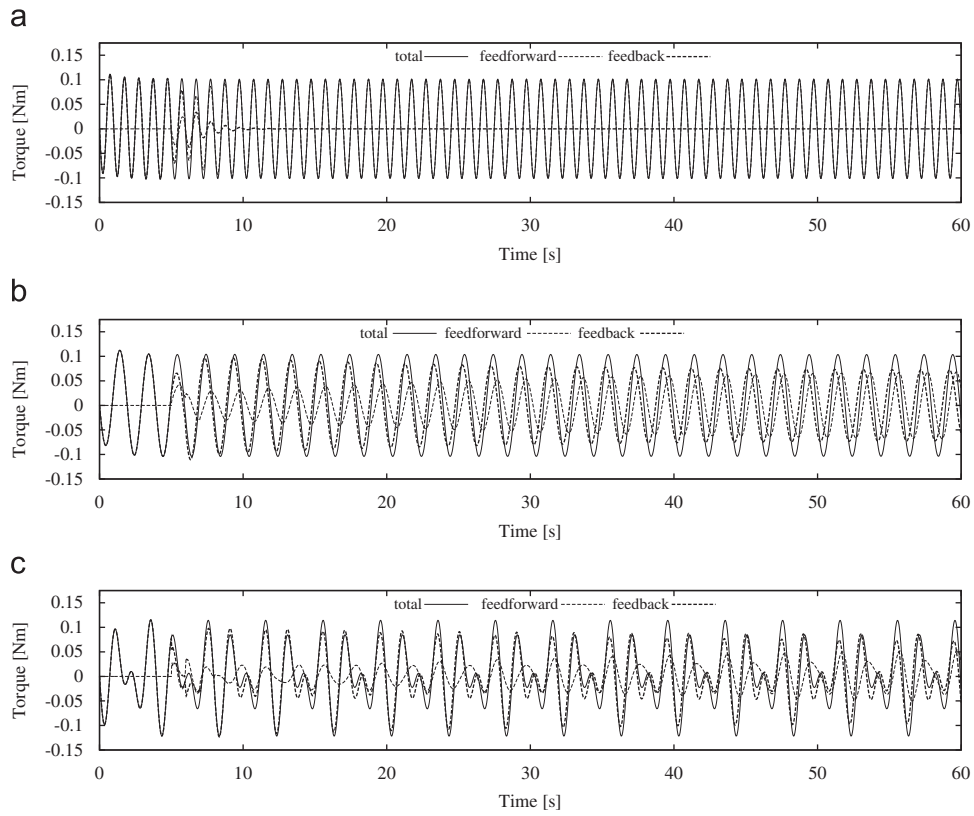
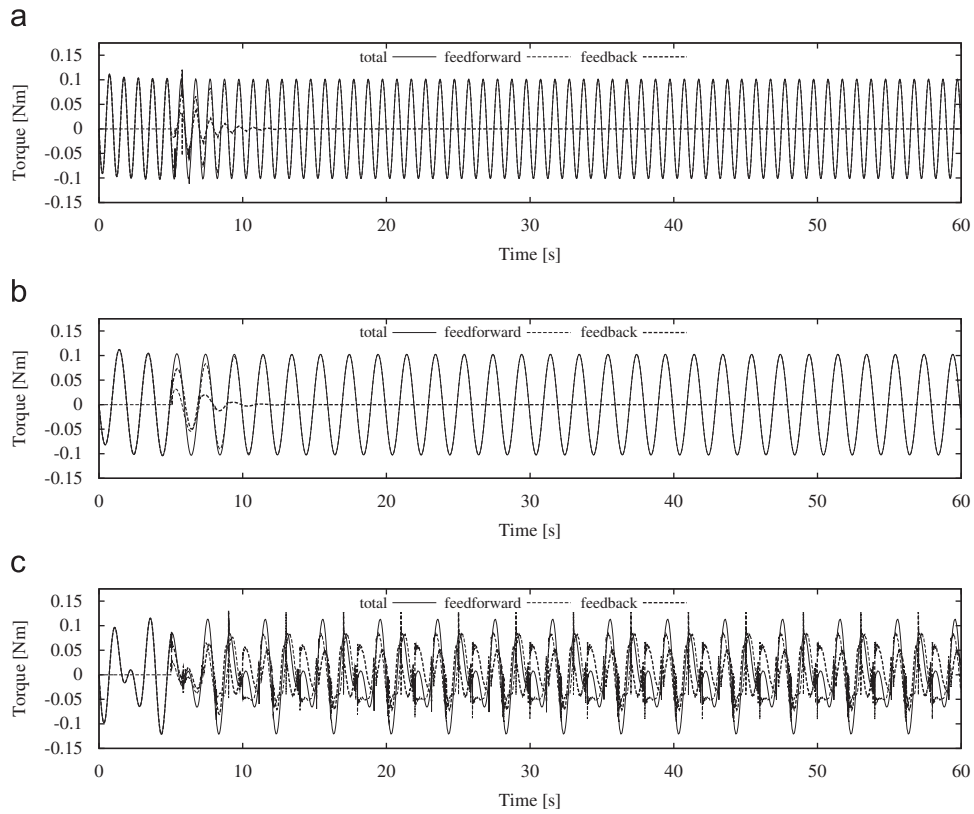**Fig. 4.** Controller without period estimation. (a–c) Torque in cases.

Fig. 5. Controller with period estimation based on the low-pass filter. (a–c) Torque in cases.

Here,

$$\langle a(t), b(t) \rangle_{T_s} = \int_{-T_s}^{0} a(t+\xi) \cdot b(t+\xi) \, d\xi. \tag{36}$$

3. This $r(t)$ should be nearly equal to 1 with every interval T_s because of its periodicity. Thus, the time $t_{max}(k)$ which gives the local maximal value beyond a threshold $r_0(<1)$ is memorized. Here, k is an integer denoting the order of the maximal value.

4. $\hat{T}_e$ is calculated by averaging some $t_{max}(k)-t_{max}(k-1)$ with every time interval T_p. Then, the estimated value of the angular frequency $\hat{\omega}_e$ is obtained by the following equation:

$$\hat{\omega}_e = 2\pi/\hat{T}_e. \tag{37}$$

5. The basic angular frequency in Fourier expansion ω is adjusted as the dynamics with the first order lag

$$\dot{\omega} = -k_{\omega}(\omega - \hat{\omega}_e) \tag{38}$$

to avoid the discontinuous changes of the angular frequency.

When the basic angular frequency is estimated, its higher harmonics are obtained by the following recurrence equation:

$$\begin{bmatrix} \cos n\omega t \\ \sin n\omega t \end{bmatrix} = \begin{bmatrix} \cos \omega t & -\sin \omega t \\ \sin \omega t & \cos \omega t \end{bmatrix} \begin{bmatrix} \cos(n-1)\omega t \\ \sin(n-1)\omega t \end{bmatrix}. \tag{39}$$

Here, let us reconsider the construction of Y_r in (14). S_k and C_k are directly related to the periodic external force: the basis functions of the expansion of the periodic external force. Y_r also contains such variables as θ, $\dot{\theta}$ that are not directly related. However, θ and $\dot{\theta}$ result in periodic if the balance is maintained under the periodic external force. This fact allows the left-hand side of (24) to be wholly expressed as the Fourier series with the basic angular frequency ω_e. Namely, Y_r is reconstructed as

$$Y_r = [\sin\omega t, \cos\omega t, \ldots, \sin n\omega t, \cos n\omega t] \tag{40}$$

and the left-hand side of (13) is expressed using (40) as well as $\hat{\phi}$ that consists of the Fourier coefficients. These definitions are adopted in the following simulations and experiments.
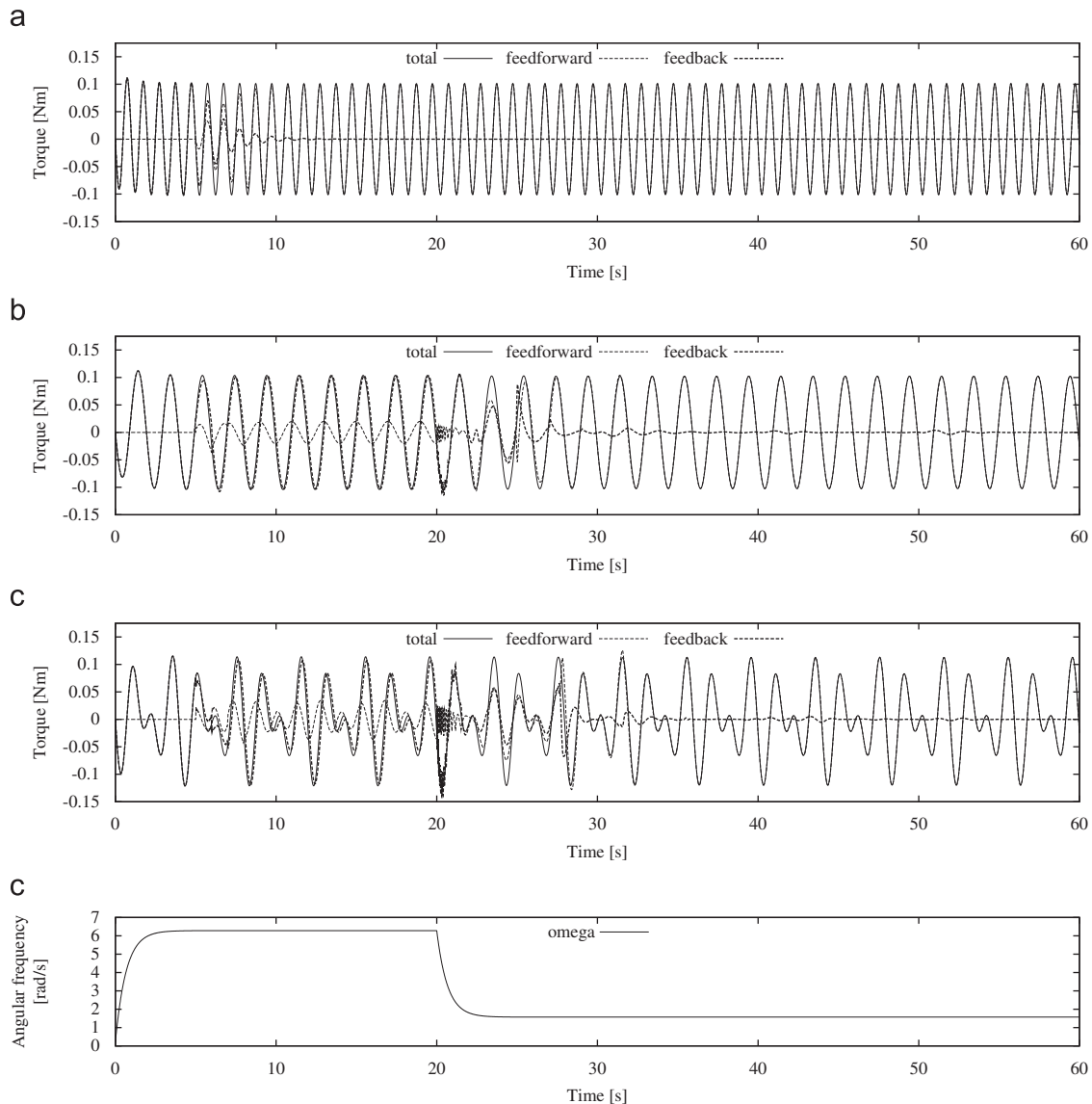

Fig. 6. Controller with period estimation using local autocorrelation. (a–c) Torque in cases and (d) estimated angular frequency in case (c).

4. Simulation

4.1. Period estimation and torque pattern learning

4.1.1. Purpose

The first simulation investigates the efficiency of the period estimation method based on the local autocorrelation, and examine whether the torque trajectory is certainly stored to the feedforward controller as the motion pattern generator.

4.1.2. Conditions

The adaptive posture change as well as the torque pattern learning was simulated for the static balance model that is mentioned in Section 2.2. The link parameters are $M = 0.50$ kg, $L = 0.2$ m, $\ell = 0.05$ m, $I = 0.025$ kg m^2. Three controllers are compared. They are applied to the same control law (12) and learning law (19), but differ in the period estimation algorithm, i.e., controller with

(i) no period estimation. ω is fixed to 1 Hz;
(ii) period estimation based on the low-pass filter (proposed in our previous paper, Ito et al., 2005). The 3-order Butterworth filter is used for the ankle joint velocity;
(iii) period estimation using local autocorrelation (proposed in this paper). The parameters are set as $T_s = 1$ s, $T_p = 10$ s, $r_0 = 0.99$ and $k_\omega = 1.5$.

In each controller, the same values are set to the following parameters: $K_d = 20$, $K_p = 70$, $K_f = 3$, $\Gamma = \text{diag}[25, \ldots, 25]$ and $n = 20$. Learning starts at 5 s. The external force is imposed by the following equation:

$$F_x = Mg\sin\alpha, \tag{41}$$

$$F_y = Mg(1 - \cos\alpha), \tag{42}$$

which is equivalent to the force that exerted on the slope with the gradient α. α is set in the following three manners:

(a) $\alpha = A\sin 2\pi f$, $A = 5°$, $f = 0.5$ Hz.
(b) $\alpha = A\sin 2\pi f$, $A = 5°$, $f = 1$ Hz.
(c) $\alpha = A_1\sin 2\pi f_1 + A_2\sin 2\pi f_2$, $A_1 = A_2 = 3°$, $f_1 = 0.5$ Hz, $f_2 = 0.75$ Hz.

The 4th-order Runge Kutta method is used with the step size 1 ms.

4.1.3. Results

The torque trajectory τ with its feedforward and feedback components τ_{ff} and $\tau_f b$ are depicted in Figs. 4, 5 and 6, respectively, for the controllers (i), (ii) and (iii). The graph is obtained for the external force in the cases (a), (b) and (c) from the top. For the controller (iii), the estimated angular frequency is shown in the bottom.

The controller (i) without the period estimation can learn the torque trajectories only if the external force has the same period as the preset value 1 Hz, as shown in Fig. 4(a). In other case, these trajectories cannot be copied to the feedforward controller completely, as shown in Fig. 4(b) and (c).

The controller (ii) can learn simple periodic trajectories with various period, as shown in Fig. 5(a) and (b). However, complex periodic trajectory that contains multiple frequency components cannot learn completely, as shown in Fig. 5(c).
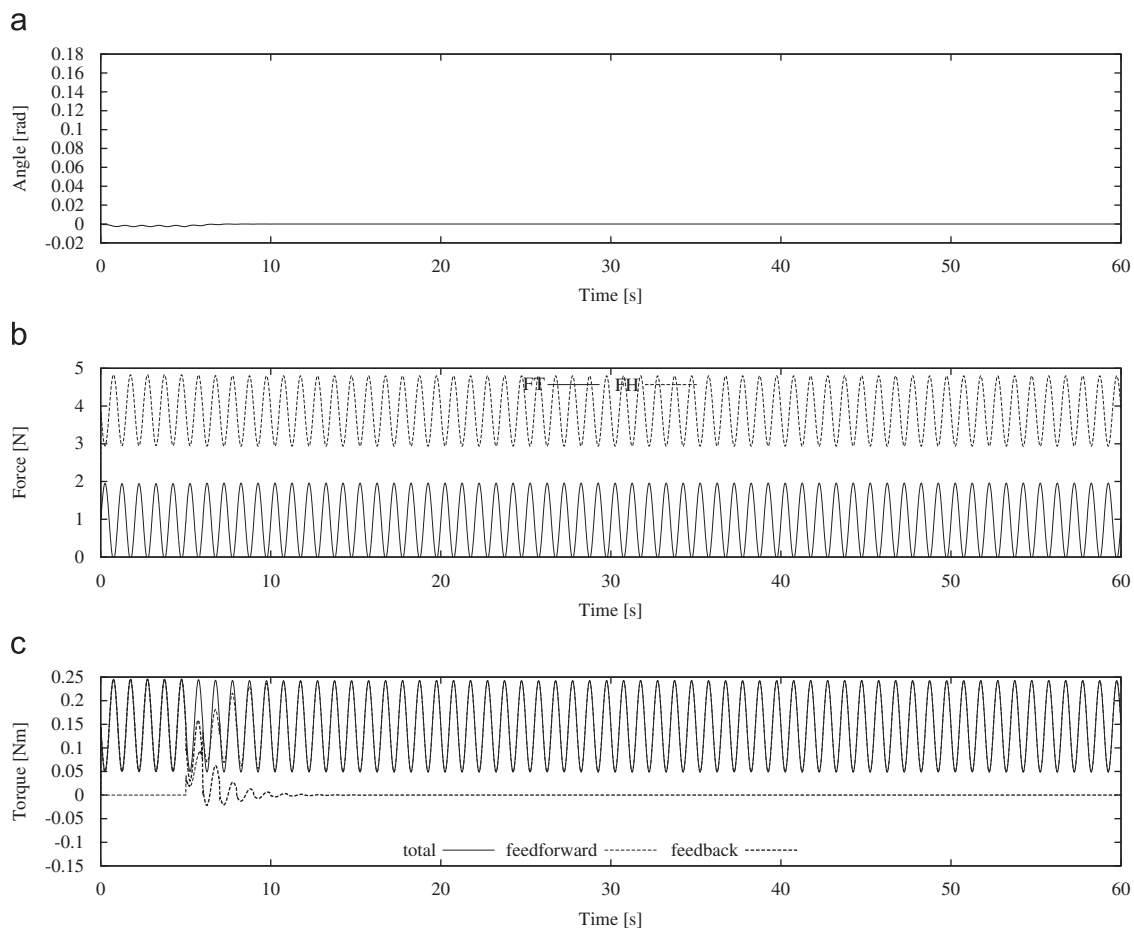

Fig. 7. Result from a PD controller. (a) Ankle joint angle. (b) Ground reaction force. (c) Torque with feedforward and feedback component.

On the contrary, the controller (iii) that is proposed in this paper can learn even complex periodic trajectories, as shown in Fig. 6(a)–(c). Fig. 6(d) demonstrates that this learning acquires the correct estimated value, i.e., about 0.25 Hz ($\pi/2$ rad/s), the greatest common divisor of $f_1 = 0.5$ Hz and $f_2 = 0.75$ Hz.

4.2. Feedback effect of ground reaction force

4.2.1. Purpose

A remarkable feature of our control and learning method is to make an effective use of the ground reaction force feedback, though many conventional methods, typically a well-known PD control, usually contain only the feedback of the joint angle deviations/velocities. Here, our method is compared to the PD control by setting zero or non-zero value to the parameter K_f, and an advantage of our method is demonstrated.

4.2.2. Conditions

A case where our method shows a noticeable advantage is that the periodic external force contains the bias component. Thus, the external force is given as (41) and (42) where the slope angle α is set as

$$\alpha = A\sin(2\pi ft) - A_{bias}, \quad A = 0.1\,\text{rad}, \quad f = 1\,\text{Hz}, \quad A_{bias} = 0.15\,\text{rad}. \quad (43)$$

The controller (iii) using the local autocorrelation to the period estimation is used with two K_f values: $K_f = 0$ as the PD control and $K_f = 0.3$ as our proposed feedback controller. The upright posture ($\theta = 0$, $\dot{\theta} = 0$) is selected as the desired posture of the PD control. Any other parameters are the same as the previous section.

4.2.3. Results

The PD control is designed to keep the ankle joint angle constant, i.e., perpendicular to the foot part, which corresponds to $\theta = 0$ in our simulation. This goal is achieved as shown in Fig. 7(a). And, the ankle joint torque is certainly learned: Feedback torque becomes zero, while the total torque is wholly composed of the feedforward component, as shown in Fig. 7(c). However, the ground reaction forces are not satisfying the unilateral condition: they must be always positive, but FT sometimes takes negative values in Fig. 7(b). It implies that the robot tumbles by the PD control in the actual experiment.

The controller containing the ground reaction force feedback, on the other hand, drives the ankle joint periodically around forward slanted position. This motion is confirmed as a sinusoidal-like trajectory around 0.15 in Fig. 8(a). Note that this offset bias is equal to A_{bias}. Thanks to this adaptive motion, the ground reaction force is always kept positive (Fig. 8(b)). Of course, the torque learning is completed as shown in Fig. 8(c).

In conclusion, our control law brings the robustness to the biped balance against the external forces. See also Ito and Kawasaki (2005).

5. Experiments

5.1. Apparatus

A simple robot, as shown in Fig. 9(a), is used to confirm the effect of the control and learning scheme proposed in this paper. It
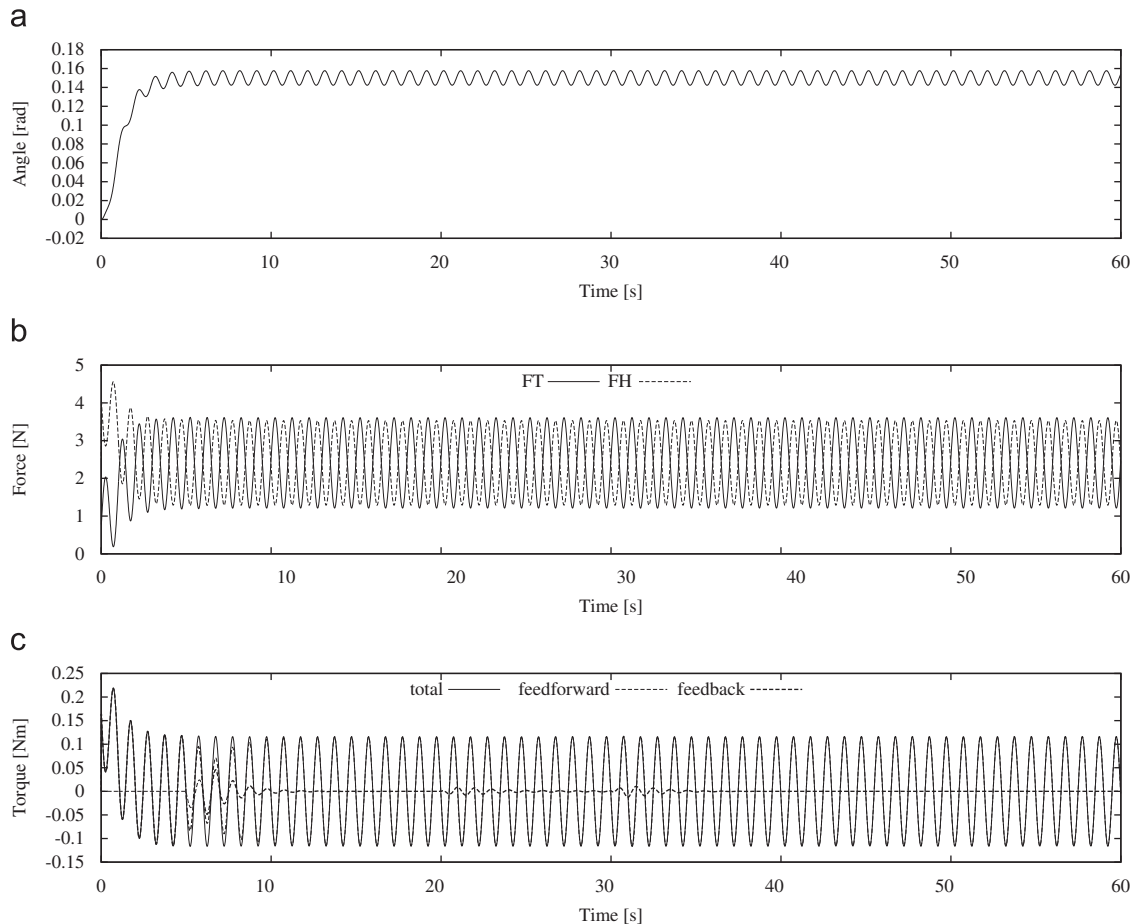
Fig. 8. Result from the controller with ground reaction force feedback. (a) Ankle joint angle. (b) Ground reaction force. (c) Torque with feedforward and feedback component.
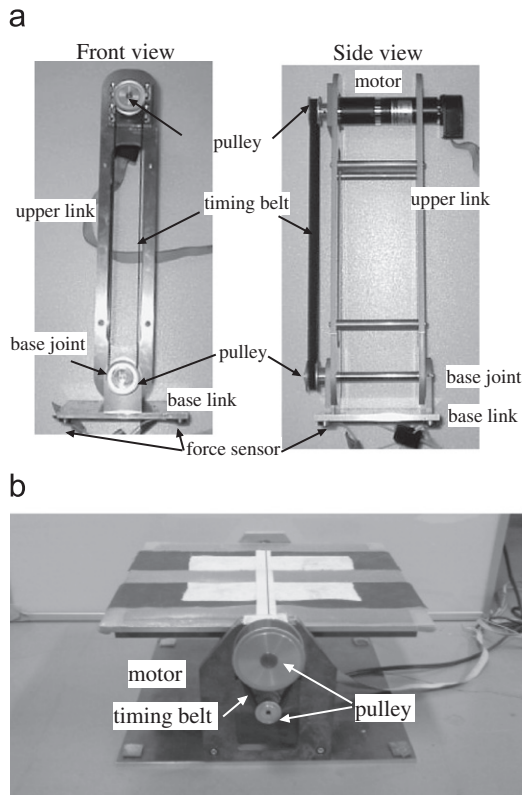
a

Front view　　　　Side view

b

Fig. 9. Apparatus. (a) A simple robot. (b) Slope stand.

has only 1° of freedom of motion at the base joint. The length of the upper link is 0.5 m, the base link is 0.1 m. The weight is 0.52 kg. The base joint at the center of the base is 0.046 m height from the ground. Four small loadcells (force sensors) are attached at each corner of the foot link. Using these loadcells, the vertical components of the ground reaction forces are measured. The angle deviation of the base joint is detected by the rotary encoder installed in the motor that drives base joint.

To impose the periodic disturbance, the slope stand as shown in Fig. 9(b) is utilized. The robot is made to balance on this periodically moving slope stand in the experiment.

The angle of the base joint as well as the slope stand is detected by the signal from the rotary encoder of the motor. This signal is input to the personal computer as the controller through the pulse counters. Their angular velocities are obtained using the digital differential filter with 0.01 rad/s cut-off frequency. Signals from the loadcells are also input through the AD converters via strain gage amplifier. And the motors are driven by the signal from the DA converter through the motor driver. The control signal is generated in every 2 ms.

5.2. Conditions

Three cases are examined. In the case (a), the period of the external force is given correctly and thus its estimation is not necessary. The external force is generated using slope stand. The slope angle α controlled using PD controller whose reference is given by the sinusoidal function

$$\alpha = A\sin2\pi f, \tag{44}$$

where $A = 0.12$ rad and $f = 0.1$ Hz. The case (b) requires the period estimation. The slope angle is changed followed by the reference trajectory (44), where the same value 0.12 rad is set to A while the different value 0.15 Hz is set to f. In the case (c), α is given by

adding up two sinusoidal waves:

$$\alpha = A_1\sin2\pi f_1 + A_2\sin2\pi f_2, \tag{45}$$

where $A_1 = 0.072$ rad $(=2.5°)$, $A_2 = 0.105$ rad $(=6°)$, $f_1 = 0.1$ Hz, $f_2 = 0.2$ Hz. The gains of the PD controller are 6.6 for proportional gain and 0.3 for the derivative gain.

The control law (12) and the learning law (19) are adopted to the controller with the following gain settings. $K_d = 3.5$, $K_p = 12.4$, $K_f = 0.1$, $\Gamma = \text{diag}[0.2, \ldots, 0.2]$, $n = 11$. The parameters for the period estimation are: $T_s = 2$ s, $T_p = 30$ s, $r_0 = 0.99$ and $k_\omega = 1.5$. Note that the period estimation is unnecessary for the case (a).

### 5.3. Results and discussions

The results of the above three conditions are shown in Fig. 10. In the case (a), a learning process in which the feedforward torque is gradually replacing the feedback torque is observed as expected. However, the feedback torque does not completely converge to zero, i.e., small vibrations remain. The main reason is a mechanical problem: the robot as well as the slope stand has small backlash in gears of motor. This backlash prevents the robot from repeating exactly the same motion, although it can be easily achieved in the simulations. The mechanical strain around the base joint may be another problem.

In the cases (b) and (c), based on the period estimation of the external force, the learning progresses so that the feedback component decreases and is gradually being replaced by the feedforward component. However, this also does not completely converge to zero. The remaining vibrations are of the same magnitude as the case (a), implying that results better than case (a) are not expected in this robot system. Accordingly, we can conclude that the torque is surely learned to the motion pattern generator, and that torque pattern generation become less dependent on the feedback signal of the ground reaction forces.

6. Discussion

This paper addresses adaptive aspects of animal motion from the viewpoint of control and learning by taking static balance as an example. The timescale in patterns of motion is focused on. On a short timescale, the motion pattern should be robust, i.e., be stabilized against sudden disturbances like environmental fluctuations. On the other hand, the motion pattern should be variable, i.e., be adjusted with respect to long-term variations such as environmental transitions. In the case of a balancing task, the former is expressed as a temporary external force, while the latter occurs as a rather permanent environmental change, such as a change of the ground slope with the standing position.

To cope with external forces with different timescales, a control and learning method is proposed: a pattern, i.e., a posture, is stabilized by the control scheme prepared for the temporary external force, while a new pattern is stored to the motion generator as a torque pattern by the learning scheme in response to the periodic disturbance. Here, information on the ground reaction force is regarded as crucial, since it is informative on the balancing situation. This is why balance is maintained based on feedback about the ground reaction forces in our control scheme. However, in the stationary condition where a periodic disturbance is exerted, the torque trajectory is learned as a motion pattern. After learning, the feedback information on the ground reaction force is no longer required: The balance can be maintained by the adaptively acquired torque trajectory in a feedforward manner. This is equivalent to a sensorless control. To achieve the sensorless control, the information on the controlled object or environment should be acquired in the controller.

a

b

c
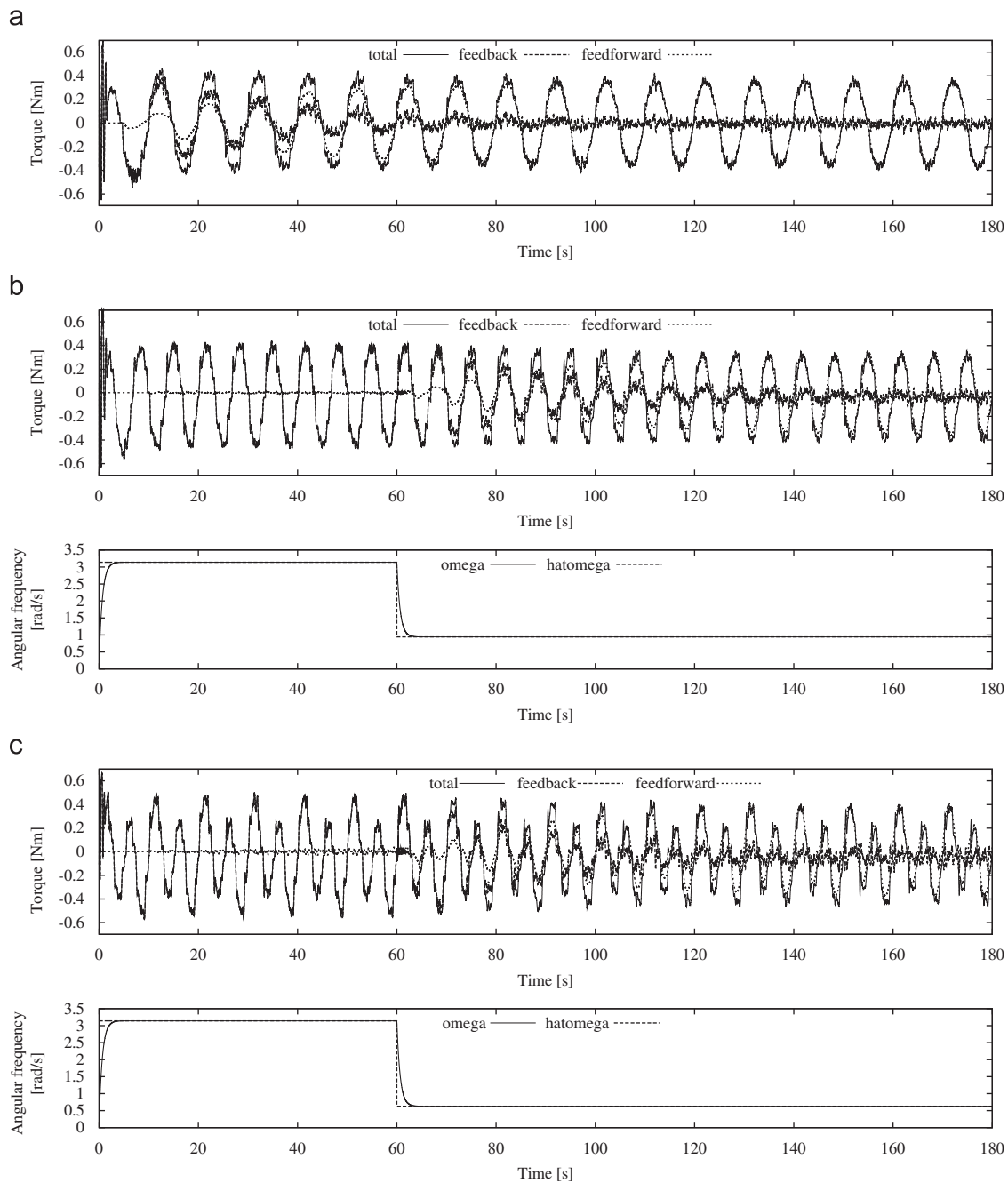
Fig. 10. Controller with period estimation based on the low-pass filter. (a–c) Experimental results for cases.

This process will be a kind of intelligence of biological systems or robots.

When non-periodic external force is exerted, the feedback controller handles them. However, stationary balancing motion never emerge from such non-periodic external force: a constant motion pattern that achieves the balance maintenance is not able to be acquired since such pattern cannot be embedded in advance in the controller if the external force is not predictable. The experimental result in Section 5 states it well: unpredictable mechanical vibration and slippage due to static frictions and backlash of gears at the joint of the robot and slope stand generate disturbances with low repeatability. Although such non-periodic component is compensated by the feedback controller, it cannot be learned, which is shown as fine fluctuations of the feedback component in Fig. 10. In summary, our control method is

constructed by combining the motion pattern generator coping with periodic excitement with a feedback controller to handle non-periodic disturbances. This dual control strategy allows us to simultaneously processes the following two tasks; the balance maintenance with respect to not only periodic but also non-periodic disturbances, and the learning of the periodic component of the control torque.

Extending a problem in our previous work (Ito et al., 2005), this paper treats the uncertainty of the period of the external force. As a solution, a period estimation based on the autocorrelation is introduced. This method requires a high-performance feedback controller that produces stable response to the periodic external forces, because the period is estimated from the periodic behavior of the ankle joint: If the ankle shows irregular behaviors, the period estimate becomes unstable or fluctuated, which resets

the learning of the motion pattern described as Fourier coefficients expanded by sinusoidal function with this estimated period. In such a case, the learning does not converge for quite a while. The high-performance feedback controller is also needed to the learning of an adequate motion pattern, since the motion pattern to be learned is a copy of the output of the feedback controller. To enhance the performance of the feedback controller, we introduced here the information of the ground reaction forces, which provides the robustness to the biped balance as shown by simulations in Section 4.2. In addition to the enhanced feedback controller, our approach brings advantages compared to the sole feedback controller in the point that the feedforward controller automatically generates an adequate motion pattern, i.e., torque trajectory independent of the sensory feedback. Furthermore, such a motion pattern is adjusted on-line collaborated with the feedback controller, although this process requires sensory information.

Finally, the ankle strategy should be discussed. The ankle strategy is well-described by an inverted pendulum model. Because of its typicality and simplicity, many studies on the human posture control adopt an inverted pendulum model from various point of view: PID control (Peterka, 2000, joint stiffness; Winter et al., 1998, intermittent control; Bottaro et al., 2008, and so on, Gatev et al., 1999; Masani et al., 2006; Qu et al., 2007). Focusing on the ground reaction forces, we also assume the ankle strategy in this paper, because disturbances by external forces are restricted only in the sagittal plane and their deviation is not so rapid in a problem example shown in Fig. 1. Actually, the frequency of the external force in experiments is not more than 0.2 Hz (Max. 8.5°). In a study of human motion measurement, the disturbances applied to the human upright posture by the moving platform is 0.5 Hz (Max. 15 cm) (Van Ooteghem et al., 2008), where the strategy gradually shifted from the ankle strategy to the multi-segment control. Thus, the ankle strategy is a valid assumption in the condition of this paper. However, some kinds of the postural control strategies are observed, and the transition among the strategy by learning is reported to actual human static balance. To extend our study, the discussions on the control and learning will be required beyond the framework of the ankle strategy.

7. Conclusion

This paper treated a biped balance control and learning when an unknown periodic external force is exerted. In this case, a human learns a motion pattern by which the biped balance is maintained with performing the stabilization task.

The above process is formulated with dynamical equations using an inverted pendulum model based on the ankle strategy. Thanks to the feedback controller based on the ground reaction force, the stationary posture adaptively changes with the external force. To store the feedback torque trajectory of this controller when a periodic external force is imposed, the motion generator is prepared in parallel. Then, the dynamics of the balancing motion is expanded to a Fourier series. The local autocorrelation of the base joint motion is efficiently used to estimate the basic angular frequency of the Fourier expansion. Next, Fourier coefficients are learned in the motion generator to replace the feedback torque trajectory by the feedforward one. In this learning, the decrement of the feedback component in the total torque output is ensured based on a Lyapunov-like lemma.

In a computer simulation, the acquisition of the torque trajectories is confirmed against unknown periodic external forces with various periods. A similar tendency toward acquiring the torque trajectories is also observed as an actual physical

phenomenon using a robot. However, the torque pattern is not completely stored, i.e., the feedback torque is not forced to zero by learning, due to some mechanical problems.

The periodicity assumption in the external forces is expected to be extended to the control of locomotion. Because stationary locomotion is regarded as periodic activity, forces disturbing the balance, such as inertial forces, also become periodic. Application to the motion planning or pattern generation of locomotion is considered as our future works.

Acknowledgements

References

Alexandrov, A., Frolov, A., Massion, J., 1998. Axial synergies during human upper trunk bending. Experimental Brain Research 118 (2), 210–220.

Behnke, S., 2006. Online trajectory generation for omnidirectional biped walking. In: Proceedings of IEEE International Conference on Robotics and Automation (ICRAf06), Orlando, FL, pp. 1597–1603.

Borghese, N., Calvi, A., 2003. Learning to maintain upright posture: What can be learned using adaptive neural network models? Adaptive Behavior 11 (1), 19–35.

Bottaro, A., Yasutake, Y., Nomura, T., Casadio, M., Morasso, P., 2008. Bounded stability of the quiet standing posture: an intermittent control model. Human Movement Science 27 (3), 473–495.

Endo, G., Morimoto, J., Matsubara, T., Nakanishi, J., Cheng, G., 2008. Learning CPG-based biped locomotion with a policy gradient method: application to a humanoid robot. The International Journal of Robotics Research 27 (2), 213–228.

Gatev, P., Thomas, S., Kepple, T., Hallett, M., 1999. Feedforward ankle strategy of balance during quiet stance in adults. Journal of Physiology 514, 915–928.

Gorce, P., Vanel, O., 1997. Behaviour synthesis of the erect stance for a biped control. Journal of Intelligent and Robotic Systems 18 (2), 127–145.

Goswami, A., 1999. Postural stability of biped robots and the foot-rotation indicator (FRI) point. The International Journal of Robotics Research 18 (6), 523–533.

Hirai, K., Hirose, M., Haikawa, Y., Takenaka, T., 1998. The development of Honda humanoid robot. In: Proceedings of 1998 IEEE International Conference on Robotics and Automation, vol. 2, pp. 1321–1326.

Hof, A., 2007. The equations of motion for a standing human reveal three mechanisms for balance. Journal of Biomechanics 40 (2), 451–457.

Horak, F., Nashner, L., 1986. Central programming of postural movements: adaptation to altered support-surface configurations. Journal of Neurophysiology 55 (6), 1369–1381.

Huang, Q., Kaneko, K., Yokoi, K., Kajita, S., Kotoku, T., Koyachi, N., Arai, H., Imamura, N., Komoriya, K., Tanie, K., 2000. Balance control of a biped robot combining off-line pattern with real-time modification. In: Proceedings of the 2000 IEEE International Conference on Robotics and Automation, vol. 4, pp. 3346–3352.

Ijspeert, A., 2008. Central pattern generators for locomotion control in animals and robots: a review. Neural Networks 21 (4), 642–653.

Ishiguro, A., Fujii, A., Hotz, P., 2003. Neuromodulated control of bipedal locomotion using a polymorphic cpg circuit. Adaptive Behavior 11 (1), 7–18.

Ito, S., Kawasaki, H., 2005. Regularity in an environment produces an internal torque pattern for biped balance control. Biological Cybernetics 92 (4), 241–251.

Ito, S., Moriki, K., Kawasaki, H., Sasaki, M., 2005. Robot experiment of torque learning for biped balance with respect to periodic external force. In: Proceedings of 12th International Conference on Advanced Robotics 2005 (ICAR'05), pp. 418–423.

Kajita, S., Tani, K., 1996. Experimental study of biped dynamic walking. IEEE Control Systems Magazine 16 (1), 13–19.

Lee, B., Kim, Y., Kim, J., 2005. Balance control of humanoid robot for hurosot. In: Proceedings of IFAC World Congress.

Lockhart, D., Ting, L., 2007. Optimal sensorimotor transformations for balance. Nature Neuroscience 10 (10), 1329–1336.

Masani, K., Vette, A., Popovic, M., 2006. Controlling balance during quiet standing: proportional and derivative controller generates preceding motor command to body sway position observed in experiments. Gait and Posture 23 (2), 164–172.

McGhee, R., Frank, A., 1968. On the stability properties of quadruped creeping gaits. Mathematical Biosciences 3 (3–4), 331–351.

Mergner, T., Schweigart, G., Fennell, L., 2009. Vestibular humanoid postural control. Journal of Physiology—Paris 103 (3–5), 178–194.

Mitobe, K., Capi, G., Nasu, Y., 2001. Control of walking robots based on manipulation of the zero moment point. Robotica 18 (06), 651–657.

Morimoto, J., Doya, K., 2001. Acquisition of stand-up behavior by a real robot using hierarchical reinforcement learning. Robotics and Autonomous Systems 36 (1), 37–51.

Nakayama, T., Kimura, H., 2004. Trajectory tracking control of robot arm by using computational models of spinal cord and cerebellum. Systems and Computers in Japan 35 (11), 1–13.

Napoleon, S., Sampei, M., 2002. Balance control analysis of humanoid robot based on ZMP feedback control. In: Proceedings of the 2002 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 2437–2442.

Nashner, L., 1976. Adapting reflexes controlling the human posture. Experimental Brain Research 26 (1), 59–72.

Nishii, J., 1999. A learning model of a periodic locomotor pattern by the central pattern generator. Adaptive Behavior 7 (2), 137–149.

Nishiwaki, K., Kagami, S., Kuniyoshi, Y., Inaba, M., Inoue, H., 2002. Online generation of humanoid walking motion based on a fast generation method of motion pattern that follows desired ZMP. In: IEEE/RSJ 2002 International Conference on Intelligent Robots and System, vol. 3, pp. 2684–2688.

Peterka, R., 2000. Postural control model interpretation of stabilogram diffusion analysis. Biological Cybernetics 82 (4), 335–343.

Prahlad, V., Dip, G., Meng-Hwee, C., 2007. Disturbance rejection by online ZMP compensation. Robotica 26 (1), 9–17.

Priplata, A., Niemi, J., Salen, M., Harry, J., Lipsitz, L., Collins, J., 2002. Noise-enhanced human balance control. Physical Review Letters 89 (23), 238101-1–238101-4.

Qu, X., Nussbaum, M., Madigan, M., 2007. A balance control model of quiet upright stance based on an optimal control strategy. Journal of Biomechanics 40 (16), 3590–3597.

Rietdyk, S., Patla, A., Winter, D., Ishac, M., Little, C., 1999. Balance recovery from medio-lateral perturbations of the upper body during standing. North American Congress on Biomechanics. Journal of Biomechanics 32 (11), 1149–1158.

Slotine, J., Li, W., et al., 1991. Applied Nonlinear Control. Prentice-Hall, Englewood Cliffs, NJ.

Sugihara, T., Nakamura, Y., Inoue, H., 2002. Real-time humanoid motion generation through ZMP manipulation based on inverted pendulum control. In: Proceedings of 2002 IEEE International Conference on Robotics and Automation, vol. 2, pp. 1404–1409.

Takanishi, A., Tochizawa, M., Karaki, H., Kato, I., 1989. Dynamic biped walking stabilized with optimal trunk and waist motion. In: Proceedings of IEEE/RSJ International Workshop on Intelligent Robots and Systems' 89, pp. 187–192.

Van Ooteghem, K., Frank, J., Allard, F., Buchanan, J., Oates, A., Horak, F., 2008. Compensatory postural adaptations during continuous, variable amplitude perturbations reveal generalized rather than sequence-specific learning. Experimental Brain Research 187 (4), 603–611.

Vukobratovic, M., Borovac, B., Surla, D., Stokic, D., 1990. Biped Locomotion: Dynamics, Stability, Control and Application. Springer-Verlag, Berlin-Heidelberg.

Winter, D., 1995. Human balance and posture control during standing and walking. Gait and Posture 3 (4), 193–214.

Winter, D., Patla, A., Prince, F., Ishac, M., Gielo-Perczak, K., 1998. Stiffness control of balance in quiet standing. Journal of Neurophysiology 80 (3), 1211–1221.

Wollherr, D., Buss, M., 2004. Posture modification for biped hu manoid robots based on Jacobian method. In: Proceedings of the 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2004), vol. 1, pp. 124–129.