A hierarchical learning model for basic locomotor patterns

Tomoko Hioki[†] [††], Jun Nishii[†]
[†] Faculty of Science, Yamaguchi University, 1677–1 Yoshida, 753–8512 Yamaguchi, Japan
[††] Department of Electronic Information, Yamaguchi Junior College, 1346–2 Daido, Hofu, 747–1232 Yamaguchi, Japan
{hioki,nishii}@bcl.sci.yamaguchi-u.ac.jp

Abstract

Most animal locomotor patterns are produced by two control systems: a higher center such as the cerebellum and the CPG (Central Pattern Generator) which is located in the spinal cord of vertebrates. We take notice of the following two questions. How are firing patterns of the CPG well coordinated? How do two motor commands from the CPG and a higher center work without confliction in order to realize a target motion? In order to answer these questions, we propose a motor learning model that a higher center learns an appropriate motor command which works as a supervisory signal for the CPG and an appropriate firing pattern of the CPG is learned by the signal. We applied this model to the learning control of a one-dimensional hopping robot. As a result, higher center and the CPG learned proper motor commands for the target motion.

Key words Central pattern generator, Learning, Motor control, Hierarchical architecture

1 Introduction

Most animal locomotor patterns, such as walking and swimming, present cyclic movements, which are produced by properly patterned firing of neurons of the CPG (Central Pattern Generator) in the central nervous system. The CPG receives feedback signals from a musculoskeletal system, however it is capable of generating motor commands even in an immobilized animal [1].

Although a firing pattern of the CPG is genetically determined at some level, learning and coordination of an adequate firing pattern would be necessary to respond to the change of body parameters due to growth and injury. However it has not been known how firing patterns of the CPG are learned. In the learning of the CPG a kind of teacher signals from a higher center to the CPG would be necessary because the CPG is situated at the lower part of a nervous system.

An experimental result with decerebrate cat indicates that smooth movements require not only the control by the CPG but also by a higher center [2]. Yanagihara et al. (1994) reported that climbing fiber responses in cerebellar vermal Purkinje cells in decerebrate cats increased during perturbed locomotion [3]. This result shows that the responses of climbing fibers represent error signals in control of movements. Therefore a higher center would be monitoring the sequence of normal gait controlled by the CPG and generating motor commands to recover from the perturbation. However it is not clear how do two motor commands from the CPG and a higher center work without confliction? To avoid this confliction, a higher center must monitor the activity of the CPG and sends a signal to control the CPG. Projections, from the brainstem of a lamprey to the CPG is observed and a stimuli to such region in the brainstem makes the lamprey take a certain postures, or swims front or back, or turns around [4]. This experimental result indicates that the neural activity of the brainstem exerts a strong influence over the activity of the CPG, and suggests that the brainstem controls the neural activity of the CPG. It has been also reported that the neural activity of the brainstem is largely affected by the firing of the CPG [5], which suggests that the brainstem is monitoring the activity of the CPG.

On the basis of these neurophysiological knowledge, we propose a hierarchical learning model of the CPG and a higher center, which solves the problems of the confliction of multiple control signals and the learning of the CPG.

2 A hierarchical learning model

Figure 1 shows a schematic representation of a hierarchical motor learning model that a higher center learns an appropriate motor command which works as a supervisory signal for the CPG to learn an appropriate firing pattern (Fig. 1). In this model, a higher center sends control signals to the CPG according to the states of the musculoskeletal system and the CPG.

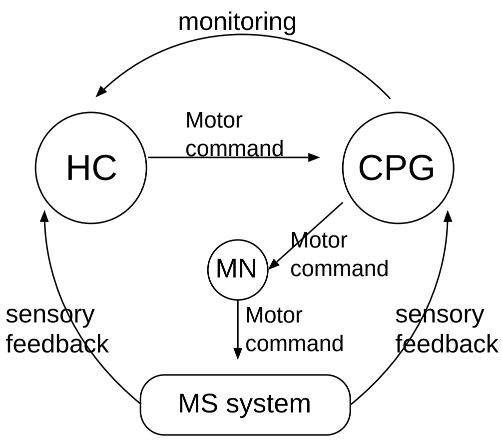

Figure 1: Schematic representation of the proposed motor control model by a higher center and the CPG. A control signal generated by a higher center is sent to a motor neuron through the CPG. HC is a higher center, MN is a motor neuron, and MS system is the musculoskeletal system.

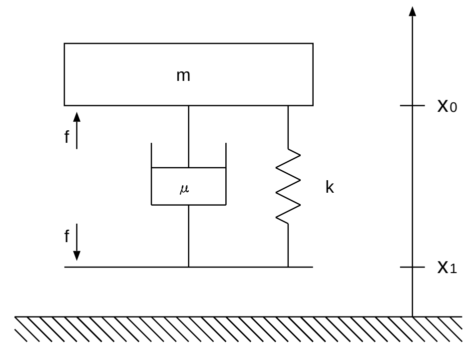

Figure 2: One-dimensional hopping robot [6]. m is the mass of the robot, x_0 and x_1 are positions of the toe and the trunk, respectively, k and μ are an elastic coefficient and a damping coefficient of the leg, respectively, and f is the force applied between the trunk and the toe when a motor neuron receives a control signal.

The control signal works as a reset signal for the CPG to cause an immediate firing which induces the firing of the motor neuron and as a suspend signal to delay the firing of the CPG, and are tuned so as to realize a target action. By sending a motor command from a higher center to motor neurons through the CPG, the state of the CPG can be controlled, by which confliction of two motor commands would be avoided.

3 Simulation method

By applying the proposed model to the learning control of a one-dimensional hopping robot, we examined the learning performance of the proposed model. The one-dimensional hopping robot consists of a trunk and a leg with a spring component and a damping component (Fig. 2 [6]). When the motor neuron receives a control signal, a force is generated between the trunk and the toe. If the force is applied to the robot in an appropriate timing in a hopping cycle, the robot can keep hopping with target height. In this simulation, the CPG is assumed to be a phase oscillator.

3.1 A learning model of its higher center

Q-learning [7] was applied to the learning of the control signals of a higher center, and the lowest height of the trunk, x_0 [m], and the phase of the CPG, θ, are used as the states for Q-learning. When the robot reaches the lowest position, an action which the higher center takes is chosen from (1) sending an excitatory signal to the CPG δ seconds later ($\delta = 0.001 \times a$ [s], $a = 0 \ldots 57$), (2) sending an inhibitory signal to the CPG at the moment, and (3) sending no signal to the CPG during a hopping cycle. Reward function is set as $10e^{-\frac{(\hat{x}_0 - x_0)^2}{0.2^2}}$. When a higher center sends a control signal to the CPG, negative reward (-5) is added to the total reward, by which it is expected that the learning of a control signal from a higher center proceeds so as that a higher center sends no signal to the CPG when the CPG sends a correct signal to the motor neuron.

3.2 A model of the CPG and a learning rule

In this simulation the phase dynamics of the CPG is assumed to take the following form:

$$\dot{\theta} = \omega + \sum_i W_i P(\theta - \varphi_i) Q(t) + P(\theta) Q_{HC}(t),$$

where θ is the phase of the oscillator, ω is the intrinsic frequency of the CPG, $P(\theta)$ is the function which shows the effect of input signals to the oscillator and is assumed to be $P(\theta) = -\sin 2\pi\theta$, φ_i indicates the phase delay of the effect, $Q(t)$ is a sensory feedback signal to the oscillators, W_i is its weight of the connection, and $Q_{HC}(t)$ is a signal from a higher center.

We use $Q(t) = \frac{v}{<|v|>}$ as the sensory feedback signal, where $<>$ means the time average (see Appendix A).

We define the signal of a higher center as $Q_{HC}(t)$=20 for action (1) in section 3.1, $Q_{HC}(t)$=-1.0 for 0.01 [s] for condition (2), and otherwise $Q_{HC}(t)$=0. The CPG sends a pulse signal to a motor neuron when the phase of the CPG becomes $\theta = 0$, which results in the force generation f =0.4 [N] for 0.15 [s] between the trunk and the toe.

We apply the learning rule proposed by Nishii(1998) [8] to the learning of the intrinsic frequency ω of the CPG and the weights W_i. The learning rule takes the form as follows:

$$\dot{\omega} = \epsilon_\omega \left\{ \sum_i W_i \langle P(\theta - \varphi_i) Q(t) \rangle + \langle P(\theta) Q_{HC}(t) \rangle \right\},$$
$$\dot{W}_i = \epsilon \langle P(\theta - \varphi_i) Q(t) \rangle \cdot \langle P(\theta) Q_{HC}(t) \rangle,$$

where ϵ and ϵ_ω are parameters determining the learning rate.

By this learning rules, the timing of the output signals of the CPG becomes in phase with the control signal from a higher center and the effect of the control signal on the phase dynamics becomes zero. Thus it is expected that only the CPG can generate an appropriate signal after learning even if the signal from a higher center is blocked.

Because continuous hopping of the robot is necessary for the learning of an appropriate control signal, we set the initial parameters of the CPG so as to enable the continuable hopping. Parameters in this simulation are shown in Appendix A.

3.3 Results

Figure 3 shows the time profile of the time averaged height of the trunk of the robot ($\bar{x}_0$) during learning. It is shown that an appropriate control signal is acquired for each target height $\hat{x}_0$ =1.1, 1.3, 1.5 [m]. Figure 4 shows the time profile of the time averaged height of robot's trunk when only the CPG sends motor control signals after learning. The robot continues hopping at desired heights, which indicates that the parameters of the CPG are tuned properly.

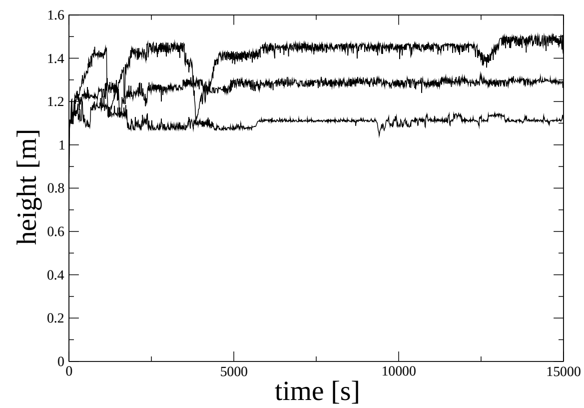

Figure 3: The time profile of the time averaged height of the trunk of the robot ($\bar{x}_0$) during learning. The target heights are set as $\hat{x}_0$=1.1, 1.3, 1.5 [m].

4 Conclusion

In our study, we proposed a hierarchical motor learning model composed of a higher center and the CPG. In this model a higher center learns an appropriate motor command, and sends it to the CPG as a supervisory signal. The CPG coordinates its natunal frequency and weights of connections with sensory feedback for an appropriate firing pattern on the basis of the signal. This model solves the problem of confliction of two motor commands, and enables the CPG to learn an appropriate control signal. In living bodies multiple control systems as the proposed model would make animal locomotor patterns robust.

By applying the proposed model to the learning control of one-dimensional hopping robot, desired hopping was successfully acquired. Although we applied the model to a simple model with single degree of freedom in this article, we expect that the model could be applied to a complicated movement such as walking and swimming.

A The parameters of the CPG during learning

The parameters of the CPG are given as follows.

$\omega = 1.0$, $W_1 = 0.2$, $W_2 = -0.4$, $\varphi_1 = 0$, $\varphi_2 = 0.8$, $\epsilon = 0.5$, $\epsilon_\omega = 0.5$.

Time average is used first-order lag low-pass filter, and time constant is 5 [s].

References

[1] S. Grillner, J. T. Buchanan, P. Wallen, and L. Brodin, *Neural Control of Rhythmic Movement in Vertebrate*, Wiley-Interscience Publication, 1988.

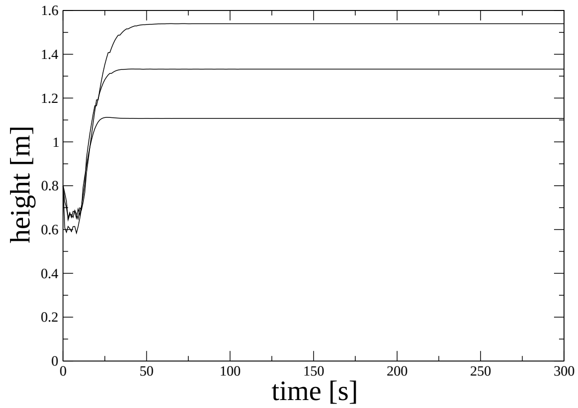

Figure 4: The time profile of the time averaged height of the trunk of the robot ($\bar{x}_0$) after learning.

[2] G. M. Shepherd, *Neurobiology*, Oxford University Press, 1994.

[3] D. Yanagihara, M. Udo, "Climbing fiber responses in cerebellar vermal Purkinje cells during perturbed locomotion in decerebrate cats", *Neuroscience Research*, Vol.19, pp.245-248, 1994.

[4] C. M. Rovainen, "Physiological and anatomical studies on large neurons of central nervous system of the sea lamprey (*Petromyzon marinus*). I. Muller and Mauthner cells", *Journal of Neurophysiology*, Vol.30, pp.1000-1023, 1967.

[5] S. Grillner, P. Wallen and L. Brodin, "Neuronal network generating locomotor behavior in lamprey: circuitry, transmitters, membrane, properties and simulation", *Annual Review of Neuroscience*, Vol.14, pp.169-199, 1991.

[6] J. Nishii , "An adaptive control model of a locomotion by the central pattern generator" *Natural to Artificial Neural Computation, Lecture Notes in Computer Science*, Vol. 930, pp. 151–157, Springer, Berlin Heidelberg, 1995.

[7] R. S. Sutton, A. G. Barto, *Reinforcement Learning: An Introduction*, MIT Press, 1998.

[8] J. Nishii, "A learning model for oscillatory networks", *Neural Networks*, Vol.11, pp.249-257 1998.