The 2010 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 18-22, 2010, Taipei, Taiwan

Emergence of bipedal walking through body/environment interactions

Shingo Shimoda, Yuki Yoshihara and Hidenori Kimura

***Abstract*— In biological regulatory systems, all computations result from spatial and temporal combination of simple and homogeneous computational media. This computational scheme realize the adaptability to unpredictable environmental changes, which is one of the most salient features of biological regulations. To investigate the learning process behind this computational scheme, we propose a learning method that embodies the features of biological systems, termed *tacit learning*. We have constructed a controller based on the notion of tacit learning and applied it to the control of the 36DOF humanoid robot to create the bipedal walking adapted to the environment. Experiments on walking showed a remarkably high adaptation capability of tacit learning in terms of gait generations, power consumption and robustness.**

I. Introduction

Adaptations to unpredictable environmental changes are the highest-priority task for all living organisms to survive in the natural environment. In the long history of evolution, living organisms have developed regulatory systems that form a network of simple and homogeneous computational media. Homogeneity here means that each computational medium follows identical activity rules and computations for regulations are carried out by accumulating local activities of computational media subject to the identical rules. In brains, for example, neurons modify their synaptic connections, quantities of releasing neurotransmitters, conductances of ion channels and so on based on their innate rules and creates behaviors adapted to the environment[1]. Similar processes of adaptations can be observed in other biological regulatory systems such as intracellular regulations[2]. We termed this feature of biological regulations *compound control*[2]–[6].

Usage of homogeneous computational media for regulations is considered to be an effective and economical way of embedding the regulatory mechanism taking into account the limited resources of biological systems[2]. One of the most serious problems of the homogeneous computational media network, however, is how to distinguish the difference of the environmental informations which are represented by homogeneous signal flows inside the network. Biological systems should have some mechanisms to distinguish them because biological systems are very sensitive to small environmental changes.

Whereas recent studies have succeeded in clarifying the details of biological computational media[7][8] and their network structures[9][10], the process of organizing the activities of the computational media and creating global behaviors still remains as an attractive research topic.

Construction of artificial controllers that are capable to adapt to unpredictable changes is the interesting approach to understand the adaptation process since artificial systems are much easier than biological systems to analyze that process[11]. Recent advances in artificial learning and adaptive methods for robot controls, however, have not reached a level of biological systems' adaptability[12]–[18]. One of the most critical problems of the conventional approaches is the way of specifying goals of learning. In many cases, goals of learning are specified in advance by using supervising signals such as teaching signals in neural networks[12]–[14] and reward functions in reinforcement learning[15], which are not changed during learning.

In the previous papers[3][4], we proposed an unsupervised learning scheme based on the features of the biological regulatory systems. We termed this learning scheme *tacit learning*. To apply tacit learning to the learning of bipedal walking, the network of the artificial computational media was implemented to the controller of 14DOF bipedal robot. The robot succeeded in learning the bipedal walking in completely model-free fashion. Balance, which is an important factor of bipedal walking, emerged within approximately 10 minutes through the motions in the real environment[19].

This success of tacit learning provides insights into the process of creating behaviors through the motions in the real environment. In that experiment, the proposed networks had some innate connections between the sensors and the motors in the form of reducing the sensor inputs by the motor actions, just like our motion of withdrawing our hands from hot plates[3]. These innate connections generated the reflexive actions through body/environment interactions. The environmental informations were taken into the network by these reflexive actions.

We consider that the way of acquiring the environmental information through the reflexive actions strongly contributes to distinguish the environmental informations in the computations using the homogeneous computational media. The process of creating behaviors from the environmental informations acquired through the reflexive actions is, however, still unclear though we experimentally succeeded in learning the bipedal walking. In the paper[3], we developed a mathematical model of the proposed network based on the individual activities of the artificial computational media. The aim of this paper is to investigate the process of creating the behavior based on the mathematical model and show the adaptability of tacit learning.

II. Tacit Learning

A. Feature of tacit learning

The fundamental idea of tacit learning comes from the feature of biological regulatory systems in which all regulations results from spatial and temporal combination of simple and homogeneous computational media. Protein-protein interactions in intracellular regulations and the networks of neurons in brains are the prominent examples of such regulatory systems.

S. Shimoda, Y. Yoshihara and H. Kimura are with RIKEN Brain Science Institute - Toyota Collaboration Center, Japan shimoda@brain.riken.jp

978-1-4244-6676-4/10/$25.00 ©2010 IEEE

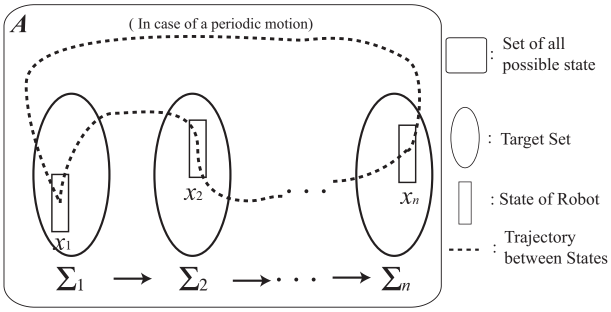

Fig. 1. Conceptual image of task and its execution

Based on this feature of biological systems, tacit learning is characterized by the following features. First, the controller for tacit learning is a network of homogeneous computational media. The learning progresses through accumulating the individual activities of the computational media that are operated by their innate rules. Second, the sensor-motor connections are organized in the network such that the sensor inputs are reduced by the motor actions. These innate connections create the reflexive actions through body/environment interactions. The combination of these reflexive actions generates the primitive motions taking the environmental information into the network. Third, no supervising signal is used in the learning process. The environmental informations taken into the network through body/environment interactions lead the primitive motions to the behaviors adapted to the environment.

B. Definition of task for tacit learning

Tacit learning should be discussed with specified target behaviors because tacit learning is the way of creating behaviors adapted to the environment achieving specified purposes. In the case of learning of bipedal walking discussed in [4], the specified target behavior was the locomotion part of walking that swings a leg forward alternately. The balance and the walking rhythm emerged through tacit learning depending on the conditions of the walking surface and the weight of the robot.

We define rigorously the way of specifying target behaviors as *task* for tacit learning. Let $\boldsymbol{x}$ denote the vector representing the state of a plant to be controlled such that the behavior is described by a transition of $\boldsymbol{x}$ from an initial state to a specified state. As an example, motions of n DOF arm are described by the transition of the state of the joint space $\boldsymbol{x} = [\theta_1\ \theta_2\ \cdots \theta_n]^T$. Here, θ_i denotes the angle of Joint i. A target behavior always becomes a transition of $\boldsymbol{x}$ from an initial state to a specified state in the sate space. These specified states are called *target states*.

State space is usually very large especially in the learning of devices with many degree of freedoms, while the target states some time involve only handful degree of freedoms. Let us take a learning of the motion of picking up an object by using a redundant arm as an example. One of the target states is the posture that the end-effector of the arm reaches the object, which is not unique for a redundant arm. In such a case, a target state can be expressed as the set of state $\boldsymbol{x}$, which we call *target set*.

The task is defined as a series of the target sets. Figure 1 illustrates the conceptual image of the definition of the task. $\boldsymbol{A}$ and $\boldsymbol{\Sigma_i}$ denote a set of all possible state of $\boldsymbol{x}$ and the target sets, respectively. In the case of bipedal walking, the target sets are defined by the motions of swing legs. The motions of supporting legs and other joints are chosen freely from the target sets. By choosing the trajectories that connect the target sets, behaviors of the plant to be controlled are created. Behaviors may be the adapted ones to the environment when appropriate trajectories are chosen. We use tacit learning to choose the trajectories for creating adapted behaviors by acquiring the environmental information through the reflexive actions as mentioned in the previous subsection.

Now, tacit learning is defined to be a process of finding adapted behaviors to the environment carrying out a specified task. In the following sections, we discuss the process of finding the behaviors by tacit learning with the experimental results.

III. Emergence of trajectory by tacit learning

A. Network structures for tacit learning

As mentioned in the previous section, tacit learning is executed through accumulating the activities of the computational media. We proposed the artificial computational media in [4] whose activities are fundamentally governed by the classical McCulloch-Pitts neuron model[20], which has the two states, firing denoted by 1 and rest by 0. The difference of our model from the original McCulloch-Pitts model is the threshold modification to keep the firing frequency in the appropriate range. The threshold of each neuron is increased $\overline{\Delta\theta}$ or decreased $\underline{\Delta\theta}$ when the neuron fires or is at rest. The mathematical expression of this neuron model is given in [3]. We call this neuron mode *Variable Threshold Neuron* (VTN).

A network of VTNs has a strong computational power such as four arithmetic operations, conditioned reflexes and input accumulation by selecting appropriate threshold patterns[4]. In the paper [3], we developed the two networks based on these computational power. One is the output regulation network, which can find the appropriate threshold pattern of VTNs to control the state of the unknown plant connected to the network to the specified reference. The other is the self-reference generation network, which can find the threshold pattern that can reduce the quantity of the input to the plant by creating the control reference through body/environment interaction. The details of these network structures are given in [3]. Here, we describe the block diagrams of these networks.

The block diagrams of the overall network configurations are described in Fig. 2. The controller C in the block diagrams is the network of VTNs called *cluster* described in Fig. 3 **a**, which is composed of a series of VTNs with a common input. The output from the cluster is the number of firing VTNs. The initial values of all thresholds in the cluster are set to be equally distributed in a single band of width $\alpha + \beta$. Here, α and β denote the values for threshold incremental step $\overline{\Delta\theta}$ and decremental step $\underline{\Delta\theta}$, respectively. Under this assumption on the thresholds, the output from the cluster becomes 0 when the input is smaller than all thresholds, and all VTNs fire when the input is larger than all thresholds. Actually, the output from the cluster becomes the saturation system described in Fig. 3 **b**. An interesting feature of the cluster activity is that the non-saturated area moves depending on the output from the cluster. Non-saturated area moves to bring the area closer to the input value as described

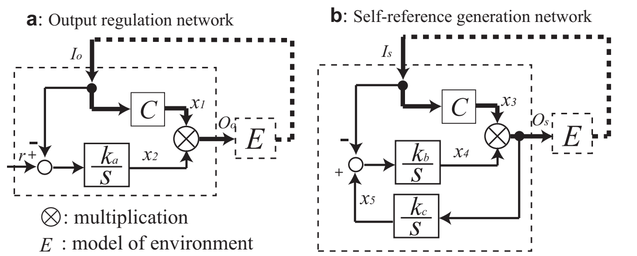

Fig. 2. Block diagrams representing output regulation network **a** and self-reference generation **b**

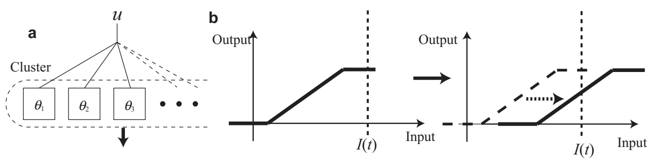

Fig. 3. Feature of cluster: **a** The cluster is a group of VTNs that have the same inputs. The output from the cluster is the number of firing VTNs. **b** The output from the cluster becomes the saturated value. The non-saturated area moves toward the input by tuning their thresholds.

in Fig. 3 **b** based on the activities of VTNs. The mathematical expression of the cluster activity is given in [3].

In the output regulation network, the output from the network O_o is the value of multiplication between the output from the cluster x_1 and the output from the integrator x_2 that is the integral value of the difference between the reference r and the input I_o as described in Fig. 2 **a**. At the equilibrium state of the output regulation network, the input I_o converges to the reference r and the cluster finds the appropriate threshold patterns in the environment. The control loop drawn by the thick lines acts as the reflexive response. The role of this loop is discussed with the experimental results in the next subsection.

In the self-reference generation network, the reference signal in the output regulation network is replaced with the integrator of the output O_s. At the equilibrium state, O_s should be 0 otherwise the output x_5 from the integrator continues changing. Thus, the self-reference generation network might be able to find the appropriate threshold pattern that tunes the output from the network to 0 if the environment allows. The mathematical expressions of these network activities are described in [3].

B. Posture control of 2DOF manipulator by tacit learning

To discuss the process of creating behaviors by tacit learning, we posed the 2DOF manipulator described in Fig. 4 **a** a task: to make a specified angle in the upper joint, called Joint 2, from the vertically standing posture as shown in Fig. 4 **b** and **c**, respectively. In this task, the angle of the lower joint, called Joint 1, was not specified but chosen freely by tacit learning. The target set $\boldsymbol{\Sigma}$ of this task is described as follows:

$$\boldsymbol{\Sigma} \quad = \quad \{(\theta_1,\ \theta_2) \mid \forall\theta_1,\ \theta_2 = \theta_d\} \tag{1}$$

which can be expressed as a line in the Cartesian space of θ_1 and θ_2 illustrated in Fig. 5. Here, θ_1, θ_2 and θ_d denote the angles of Joints 1 and 2, and the desired angle of Joint 2, respectively.

We used the self-reference generation network to control Joint 1 and the output regulation network to Joint 2 as

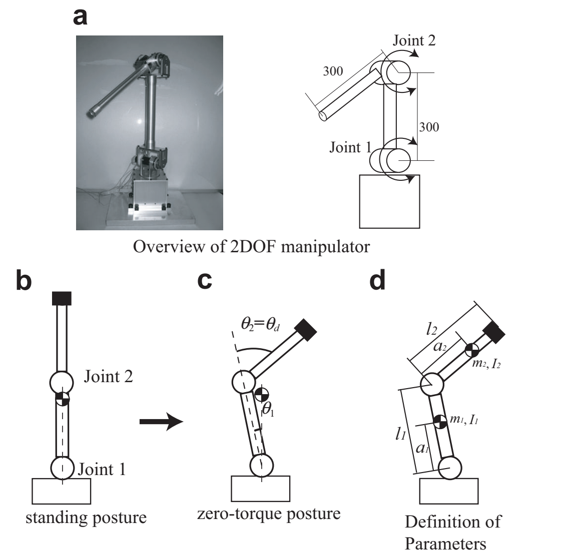

Fig. 4. 2DOF manipulator:**a** Overview of 2DOF manipulator. **b** Posture that the manipulator is standing vertically. **c** The upper joint (Joint 2) is specified to θ_d and the lower joint (Joint 1) is remained freely to choose the angle to adapt to the environment. The posture that joint 1 is controlled to be the center of mass of the manipulator is on the vertical line passing through the attachment point of Joint 1 is called zero-torque posture. **d** Definition of parameters

described in Fig. 6. In this configuration, the plant can be described as follows:

$$\boldsymbol{M}\begin{bmatrix} \ddot{\theta}_1 \\ \ddot{\theta}_2 \end{bmatrix} + \boldsymbol{B} = \boldsymbol{U} \tag{2}$$

$$\boldsymbol{Y} = \boldsymbol{C}\boldsymbol{\theta} \tag{3}$$

$$\boldsymbol{M} = \begin{bmatrix} I_1 + m_1 a_1^2 + l_1^2 m_2 + \zeta + 2\xi\cos\theta_2 & \zeta + \xi\cos\theta_2 \\ \zeta + \xi\cos\theta_2 & \zeta \end{bmatrix}$$

$$\boldsymbol{B} = \begin{bmatrix} -\xi(2\dot{\theta}_1 + \dot{\theta}_2)\dot{\theta}_2\sin\theta_2 + k_1\cos\theta_1 + k_2\cos(\theta_1+\theta_2) \\ \xi\dot{\theta}_1^2\sin\theta_2 + k_2\cos(\theta_1+\theta_2) \end{bmatrix}$$

$$\boldsymbol{\theta} = \begin{bmatrix} \theta_1 & \dot{\theta}_1 & \theta_2 & \dot{\theta}_2 \end{bmatrix}^T,\quad \boldsymbol{U} = \begin{bmatrix} u_1 & u_2 \end{bmatrix}^T,$$

$$\boldsymbol{Y} = \begin{bmatrix} y_1 & y_2 \end{bmatrix}^T,\quad \boldsymbol{C} = \begin{bmatrix} 1.0 & 0.1 & 0.0 & 0.0 \\ 0.0 & 0.0 & 1.0 & 0.1 \end{bmatrix},$$

$$\zeta = I_2 + m_2 a_2^2,\quad \xi = l_1 m_2 a_2,$$

$$k_1 = (m_1 a_1 + m_2 l_1)g,\quad k_2 = m_2 a_2 g$$

I_i : Inertia moment of each arm (See Fig.4)
l_i : Length of arm
a_i : Length from joint to center of gravity
m_i : Mass of arm (m_2 includes mass of payload.)

The most interesting feature of this experiment is how the self-reference generation network for Joint 1 finds its values during the learning. Fig. 7 illustrates the experimental result. A payload with weight 360g was attached to the top of the manipulator in this experiment. The angle of Joint 2 smoothly converged to the pre-defined reference, which is $\pi/4$ rad in this experiment. Joint 1 was rotated in the positive direction at first because the balance of the manipulator

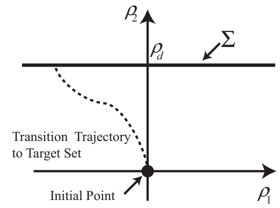

Fig. 5. Representation of target set $\mathbf{\Sigma}$ in Cartesian space: The state in $\mathbf{\Sigma}$ and the trajectory to the state are chosen through body/environment interaction.

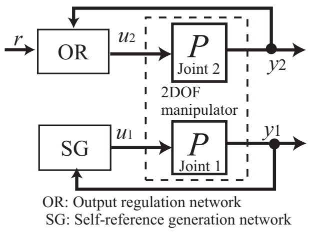

Fig. 6. Controller configuration for 2DOF manipulator control

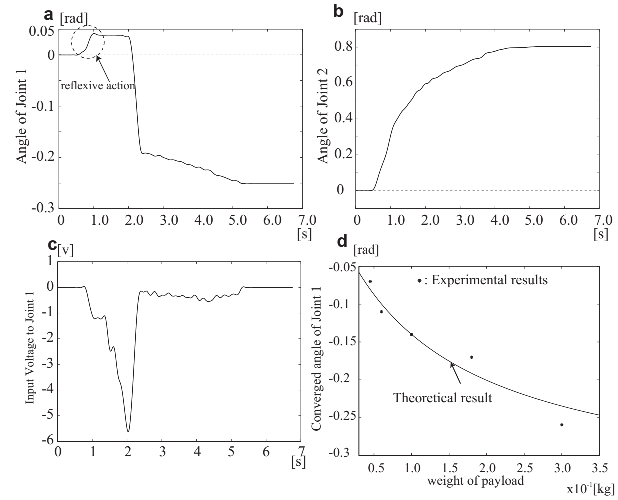

Fig. 7. Experimental results:**a** and **b** describe the time histories of Joint 1 and Joint 2, respectively. **c** is the time history of the voltage working on Joint 1. The voltage was eventually converged to 0. **d** describes the difference of converged angle of Joint 1 when the weight of the payload was changed.

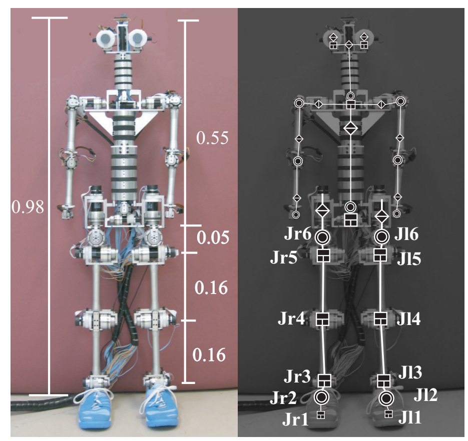

Fig. 8. Overview of 34DOF humanoid robot

was broken by the motion of Joint 2. The reflexive action that was mainly controlled by the loop described by thick lines in Fig. 2 **b** appeared immediately after Joint 1 was rotated by the loss of balance. The reflexive action of Joint 1 stimulated other network loops in Fig. 2 **b**. The angle of Joint 1 finally converged to -0.26rad, where the manipulator can keep balancing without torque to Joint 1 as shown in Fig. 7 **c**. Through tacit learning, the self-reference generation network worked to find *zero-torque posture*, for which the torque to Joint 1 is zero.

The equilibrium point of the self-reference generation network is described as follows:

$$y_1(t) = x_5(t),\ x_4(t) = 0.0,\ u_1(t) = 0.0,\ x_3(t) = \frac{N\beta}{\alpha+\beta} \quad (4)$$

Here, N denote the number of VTNs in the cluster. Under the mechanical limitation of Joint 1 which is $-\pi < \theta_1 < \pi$, the angle of Joint 1 at the equilibrium point becomes

$$\theta_1 = \tan^{-1}\left(\frac{k_2\cos\theta_d + k_1}{k_2\sin\theta_d}\right), \quad (5)$$

which corresponds to zero-torque posture. The reflexive action led the angle of Joint 1 to the above equilibrium angle that was decided by the environment, the body parameters of the manipulator and the desired angle of Joint 2 without any explicit information about zero-torque posture.

The reflexive action of Joint 1 was created by body/environment interaction, which was the loss of balance by attracting the gravitational force to the manipulator body in this case. Thus, when the body parameters were changed, the motion of Joint 1 was automatically changed. $*$ in Fig. 7 **d** describes the converged angle of Joint 1 when the weight of payload was changed. Without information to the network about the changes, the angle of Joint 1 converged to the neighborhood of the equilibrium point calculated through Eq. (4) that is illustrated in Fig. 7 **d** as the solid line.

IV. Emergence of bipedal walking by tacit learning

A. Definition of task for bipedal walking

We apply tacit learning to a emergence of bipedal using 36DOF humanoid robot described in Fig. 8. As discussed in Section II, the motions of putting the legs forward alternately were used as the target sets. Actually, we take four postures in one step as the target sets as described in Table I. In the target sets, we specified the motion of the swing legs and don't care the motion of supporting leg like Joint 1 in 2DOF manipulator experiments. The output regulation networks were used to control the specified joints and the self-reference regulation networks were used for other joints.

B. Experiments of bipedal walking

The reference values of the specified joints in the experiments are summarized in Table I. The target set was switched to the next one when the specified angles converged to the references. To create the periodic motion, Σ_8 and Σ_1 were connected.

The movies of the experiments are on [19]. At the initial state, the robot fell down even though the legs moved forward. After about 10 minutes, the motion of the supporting leg was tuned and the robot kept walking. Fig. 9 describes the lateral angle of the hip joint (Jr 6 in Fig. 8) before and after learning the walking. The motion of the joint that rotated randomly in the initial couple of minutes became periodic

TABLE I

TARGET POSTURE FOR WALKING AND SPECIFIED ANGLES

Target set (key frame)	Description	Specified DOF (Value in experiment [rad])
Σ_1	Balance on Right Leg	Jr 6 (0.08) Jl6 (-0.08)
Σ_2	Left Leg Up	Jl4 (0.4) Jl5 (0.2)
Σ_3	Left Leg Down	Jl4 (0.0) Jl5 (0.2)
Σ_4	Waiting after Left Leg Step	-
Σ_5	Balance on Left Leg	Jr6 (-0.08) Jl6 (0.08)
Σ_6	Right Leg Up	Jr4 (-0.4) Jr5 (-0.2)
Σ_7	Right Leg Down	Jr4 (-0.0) Jr5 (-0.2)
Σ_8	Waiting after Right Leg Step	-

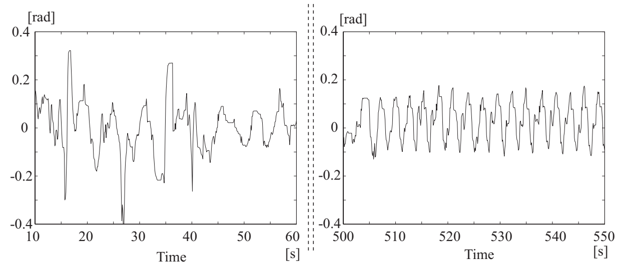

Fig. 9. Time histories of the joint Jr6

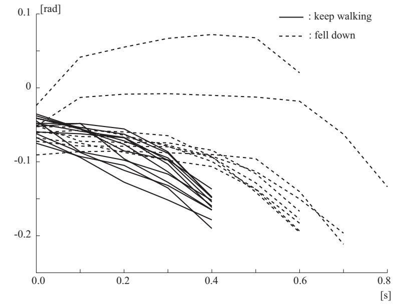

Fig. 10. The trajectories of the joint Jr6 when the state moves from $\mathbf{\Sigma_2}$ to $\mathbf{\Sigma_3}$

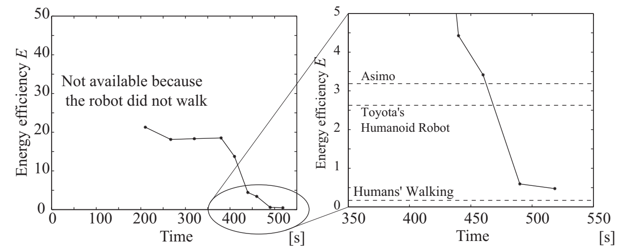

Fig. 11. Time history of energy efficiency E described in Eq. (6) during learning bipedal walking

gradually and eventually periodic motion emerged after 10 minutes.

Fig. 10 describes the trajectories of Jr 6 when the state moved from $\mathbf{\Sigma_2}$ to $\mathbf{\Sigma_3}$. The broken lines represent the trajectories that were used before learning was complete. These lines appeared at the first 3 minutes. The solid lines are the trajectories after the robot became able to walk continuously, which appeared in the final 2 minutes. The trajectory modification from the broken lines to the solid lines happed in the process of searching the equilibrium point by the reflexive actions, which was the same process of Joint 1 in 2DOF manipulator experiments discussed in the previous section. We observed the similar convergences of the trajectories of other unspecified joints.

C. Adaptability of bipedal walking to environment

Our interest is how well the created walking gait was adapted to the environment. We discuss on this problem in the following three aspects.

The first is the efficiency of the walking gait. The efficiency is one of the important indexes to discuss adaptability of walking gaits to the environment[21]. It is natural to think that the better efficiency implies the more adapted gait to the environment. We use the following index to evaluate the efficiency[22]:

$$E = \frac{\text{energy comsumption}}{(\text{mass of the robot}) \times (\text{traveled distance})}. \qquad (6)$$

Fig. 11 describes the time history of this index during learning the walking. The efficiency of human's walking and other full-control humanoid robot walkings are also illustrated in Fig. 11 as the broken lines[22]. The result shows that the efficiency of our robot was getting better as the learning progressed. At the final stage of learning, it became less than one-fifth of other full-controlled humanoid robots and almost the same value of human's walking. This remarkably high efficiency level was achieved by the reduction of the power consumption by keeping balance without torque during walking. This walking style is similar to the humans' walking and corresponds to the zero-torque posture in 2DOF manipulator experiments.

The second is on the autonomous changes of the walking rhythm. In the experiment, we did not set any time dependent parameters. The periodic walking rhythm emerged through body/environment interaction. Thus, when the body parameters and/or the environment were changed, the rhythm was automatically changed as described in Fig. 12 in which the weight of the robot was changed abruptly after learning the walking. Without any explicit information about the weight change, the rhythm was tuned slower when the weight became heavier and *vice versa*. These changes are reasonable to adapt the behaviors to the weight changes.

The final one is on the difference of the torque working on the ankle depending on the walking surfaces. As you can see in the movies on [19], our method succeeded in creating the walking on natural turf, not only on the flat and hard surface in lab. Fig. 13 **a** describes the time histories of the torque working on the ankle after learning the walking. The results show that the controller required much larger torque for the walking on turf than that in lab. We can observe the similar changes of torque in our walking. Fig. 13 **b** describes the electromyogram (EMG) data of the tibialis anterior muscle which controls the angle of the ankle when a human walked on asphalt and on sand beach. As seen in Fig. 13 **b**, we know empirically that we use much more power to keep balance while walking on sand beach than on asphalt. Thus, the emergence of appropriate torque depending on the surface conditions in our experiments implies that body/environment interactions created the walking gate adapted to the environment.

The observations on the above three aspects imply that the behaviors created by tacit learning are not just adapted to the environment, but share many features with biological behaviors. We believe that these similarities come from the

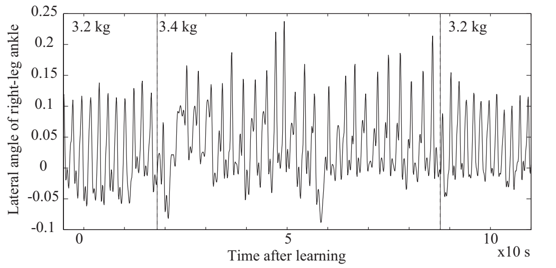

Fig. 12. Changes in walking rhythm depending on robot weight

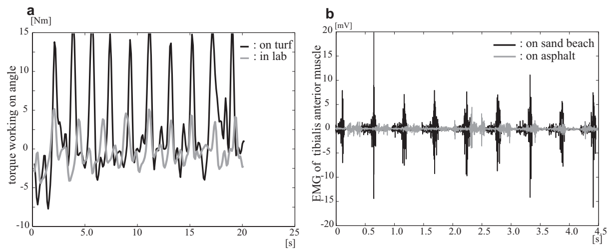

Fig. 13. Difference in torque working on ankle depending of difference of walking surface

feature of the activity rules of VTNs and their network, which change their thresholds to reduce their outputs when the environmental inputs increased. In tacit learning, the environmental informations were analyzed based on this feature. These processes tend to save energies in carrying out a specified task in the environment. This would also be important for biological systems to increase the chance of survival in the natural environment.

V. Conclusion

The notion of tacit learning has been introduced to develop the artificial control system with remarkable adaptability to unpredictable environmental changes. The fundamental computational algorithm of tacit learning is based on the feature of biological regulatory systems in which all regulations result from spatial and temporal integration of homogeneous computational media that act subject to innate rules. A network of the homogeneous computational media that connects the sensors and the motors in a proper ways is of great advantage to orchestrate the flow of heterogeneous environmental informations. We developed the networks of the artificial computational media and implemented them to control the humanoid robot.

The experimental results showed that the reflexive actions originated with the innate sensor-motor connections in the network led the primitive motions to the sophisticated behaviors adapted to the environment. Even small changes of the environment influenced the learning results because the reflexive actions were caused by body/environment interactions. The three observations discussed in Section IV verified the high adaptability of tacit learning in the natural environment.

In the proposed networks, the environmental informations taken into the network by the reflexive actions played the roles of supervising signals for learning. This learning scheme is strongly associated with the notion of *affordance*[23] that is recognized as the key factor in cognitive and intelligence. In our case, the environmental informations were mainly used to create the motions of the joints without the concrete references and led to the adapted behaviors. The creation of the meaningful behaviors from the purposeless actions by using the environmental informations should be the essential process to establish adaptation and intelligence in man-made machines.

Acknowledgements

This work was done under the full support of Toyota Motor Co. We are very grateful for their technical and financial assistance.

References

[1] M. F. Bear, B. W. Connors, and M. A. Paradios, "Neuroscience: Exploring the brain (third edition)," *Lippincott Williams and Wilkinse*, 2006.
[2] R. J. Tanaka and H. Kimura, "Mathematical classification of regulatory logics for compound environmental changes," *Journal of Theoretical Biology*, vol. 251, pp. 363–379, 2008.
[3] S. Shimoda and H. Kimura, "Bio-mimetic Approach to Tacit Learning based on Compound Control," *IEEE Transactions on Systems, Man, and Cybernetics- Part B*, vol. 40, no. 1, pp. 77–90, 2010.
[4] S. Shimoda and H .Kimura, "Neural Computation Scheme of Compound Control: Tacit Learning for Bipedal Locomotion," *SICE Journal of Control, Measurement, and System Integration*, vol. 1, no. 4, pp. 275–283, 2008.
[5] R. J. Tanaka, H. Okano, and H. Kimura, "Mathematical description of gene regulatory units," *Biophysical Journal*, vol. 91, pp. 1235–1247, 2006.
[6] H. Kimura, H. Okano, and R. J. Tanaka, "Stochastic approach to intracellular regulation and computational theory of biological control," *Journal of Theoretical Biology*, vol. 248, pp. 590–607, 2007.
[7] A. L. Hodgkin and A. F. Huxley, "Currents carried by sodium and potassium ions through the membrane of the giant axon of *loligo*," *Journal of physiology*, vol. 116, pp. 449–472, 1952.
[8] H. Markram and M. Tsodyks, "Redistribution of synaptic efficacy between neocortical pyramidal neurons," *Nature*, vol. 382, pp. 807–810, 1996.
[9] G. M. Shepherd, "The synaptic organization of the brain," *Oxford University Press*, 2003.
[10] F. Jacob and J. Monod, "Genetic regulatory mechanisms in the synthesis of proteins," *Journal of Molecular Biology*, vol. 3, pp. 318–356, 1961.
[11] M. Kawato, K. Furukawa, and R. Suzuki, "A hierarchical network model for motor control and learning of voluntary movement," *Biological Cybernetics*, vol. 57, pp. 169–185, 1987.
[12] M. L. Minsky and S. A. Papert, "Perceptron," *MIT Press*, 1969.
[13] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323-9, 1986.
[14] Y. Kuniyoshi, Y. Yorozu, S. Suzuki, S. Sangawa, Y. Ohmura, K. Terada, and A. Nagakubo, "Emergence and development of embodied cognition: A constructivist approach using robots," *Progress in Brain Research*, vol. 164, pp. 425–445, 2007.
[15] A. G. Barto, R. S. Sutton, and C. W. Anderson, "Neuronlike adaptive elements that can solve difficult learning control problems," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 13, pp. 834–846, 1983.
[16] K. J. Astrom and B. Wittenmark, "Adaptive control," *Addison Wesley*, 1989.
[17] J. E. Slotine and W. Li, "Applied Nonlinear Control," *Prentice Hall*, 1991.
[18] R. A. Brooks, "A robust layered control system for a mobile robot," *IEEE journal of Robotics and Automation*, vol. 2, pp. 12–23, 1086.
[19] "http://btcc.nagoya.riken.jp/biologic/movies.html," .
[20] W. S. McCulloch and W. Pitts, "A logical calculus of the idea immanent in nervous activity," *Bull. Math. Biophys.*, vol. 5, pp. 115–133, 1943.
[21] A. L. Hodgkin and A. F. Huxley, "Chimpanzee locomotor energetics and the origin of human bipedalism," *PNAS*, vol. 104, no. 30, pp. 12265–12269, 2007.
[22] S. Collins, A. Ruina, R. Tedrake, and M. Wisse, "Efficient bipedal robots based on passive-dynamic walkers," *Science*, vol. 307, pp. 1082–1085, 2005.
[23] J. J. Gibson, "The ecological approach to visual perception (new edition)," *Psychology Press*, 1986.