2013 IEEE/RSJ International Conference on
Intelligent Robots and Systems (IROS)
November 3-7, 2013. Tokyo, Japan

Locomotion Selection Strategy for Multi-Locomotion Robot based on Stability and Efficiency

Taisuke Kobayashi[1], Tadayoshi Aoyama[2],
Masafumi Sobajima[1], Kosuke Sekiyama[1], and Toshio Fukuda[3]

***Abstract*— This paper shows improvement of stability and efficiency for mobility using locomotion selection strategy. First strategy is the selection of a gait relying on locomotion rewards. The locomotion reward has been proposed as an indicator for selection algorithm based on Falling Risk and the moving speed. This strategy has achieved a capability of large changes of uncertainties, such as a steep slope. Second strategy is adjustment of moving speed by the extended locomotion reward that explicitly shows the relationship between the moving speed and Falling Risk. The robot aims at the maximum moving speed without a falling, and removes small changes of uncertainties as a result. We performed an experiment in order to confirm effects of two strategies in an environment that includes a rough terrain as a small uncertainty and two steps as a large uncertainty. The robot improved the moving speed about 37.5% from the case of only using the gait selection strategy.**

I. INTRODUCTION

Robots aimed at working in our living space have been developed, such as entertainment robots and home robots. High mobility is required for these robots in order to adapt to various terrains in human living environment such as stairs and slopes. In this research, mobility is defined in terms of efficiency and stability. While many criteria to characterize efficiency can be raised from viewpoints of energy, locomotion distance, task complexity etc., we consider moving speed as efficiency. In general, high-speed walking would weaken the stability due to the influence of modeling errors and tracking delay of the robot. On the other hand, low-speed walking has better stability since the walking model is close to static walking. That is, these two factors are in a trade-off relation under the same body. The robot is required to balance between stability and efficiency while walking. Then, how to balance is essential to improve total mobility.

We have dealt with such a trade-off problem throw Multi-Locomotion Robot (MLR) [1]. MLR achieves high mobility

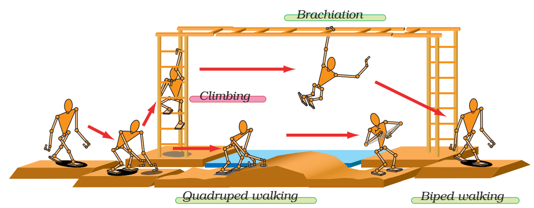


Fig. 1. Concept of Multi-Locomotion Robot

by selecting the most appropriate gait depending on its surrounding as shown in Fig. 1. We have realized the multiple gaits, such as bipedal walking with high efficiency, quadruped walking having high stability, and transition motion between the gaits. Our remaining challenge is an autonomous selection of the gait.

In previous research on the selection algorithm, two approaches have been studied. First approach is based on external environment where the robot goes through a narrow space by controlling upper body posture [2]. It is however difficult to deal with dynamic environment and internal error.

Second one used internal error information. Using predictive value in the future of the ZMP error, Ogata et al. achieved the selection of shock absorbing behavior and falling avoidance behavior when the robot was pushed in upright [3]. Toda et al. proposed a sensor-based gait generation method [4]. To estimate the road surface from the angular momentum and ZMP, the robot selects the static gait with high stability or the inverted pendulum model gait with high efficiency depending on the estimated road surface. Renner et al. realized stabilization by stopping when the robot received the disturbance during walking, depending on the degree of instability caused by the error of angle and angular velocity of the robot posture [5]. A drawback in these methods is difficulty in coping with rough terrain like a steep slope.

Considering these problems of previous work, we have proposed novel locomotion selection algorithm by Falling Risk and moving efficiency [6], [7]. Falling Risk is defined as an indicator of uncertainties using Bayesian Network in order to evaluate the state of the robot easily. The robot determines locomotion reward based on Falling Risk and moving speed against each gait, and selects the gait with the maximum locomotion reward. As a result, the robot can move in the environment that is difficult to travel by single locomotion, maintaining the maximum efficiency.

However, several issues remain to be solved. Relatively small changes of the state may bring frequent transitions

[1]T. Kobayashi, M. Sobajima and K. Sekiyama are with the Department of Micro-Nano Systems Engineering, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, 464–8603, JAPAN kobayashi@robo.mein.nagoya-u.ac.jp sobajima@robo.mein.nagoya-u.ac.jp sekiyama@mein.nagoya-u.ac.jp fukuda@mein.nagoya-u.ac.jp

[2]T. Aoyama is with the Department of System Cybernetetics, Hiroshima University, 1-4-1, Kagamiyama, Higashi Hiroshima, 739-8527, JAPAN aoyama@robotics.hiroshima-u.ac.jp

[3]T. Fukuda is with the Institute for Advanced Research, Nagoya University, Furo-cho-1,Chikusa-ku, Nagoya, 464-8603, JAPAN, the Faculty of Science and Engineering, Meijo University, JAPAN, and the Intelligent Robotics Institute, School of Mechatronic Engineering, Beijing Institute of Technology, CHINA. fukuda@mein.nagoya-u.ac.jp

978-1-4673-6357-0/13/$31.00 ©2013 IEEE

of gait, and it will result in a loss of moving efficiency. As a solution for the selection of two factors under such a small uncertainty, adjustment of moving speed is valid. The researches to control moving speed of walking have been done for a long time. These researches show that the robot cannot determine speed explicitly since speed is determined by the interaction with the environment such as CPG walking [8], [9]. Other researches do not take into account the determination of the optimal speed, only stated a method of control to the target speed [10].

In this study, we propose two strategies using selection algorithm and pursuit the maximum moving efficiency without falling. First strategy is the selection of a gait we have already proposed [7]. Second one is the adjustment of the moving speed. When the state is stable, moving speed is increased in order that the robot should have priority over the moving efficiency. During unstable state such as on a rough terrain, the robot is expected to stabilize by reducing the moving speed to improve a tracking accuracy of the COG. In addition, we want to integrate two strategies based on the same locomotion reward. We aim to determine the optimal speed consciously and autonomously, extending the locomotion selection algorithm in order to deal as a selection problem. By this extending, both strategies are based on the same locomotion reward although each strategy works on a degree of uncertainties suitable for each capability. For validation of proposed two strategies, we perform an experiment in an environment that includes a rough terrain as a small uncertainty and two steps as a large uncertainty. As an experimental result, the robot traveled the complex environment where the robot cannot go through only using biped walking. The moving speed improved about 37.5% against the case that the robot travels only using first strategy.

II. MULTI-LOCOMOTION ROBOT

A. *Gorilla Robot* III

MLR is a new bio-mimetic robot that can perform in standalone several kinds of locomotion such as biped walking, quadruped walking, and brachiation [1]. We developed Gorilla Robot III as a prototype of MLR: its height is about 1.0 [m] and weight is about 24.0 [kg]. An overview and the link structure of Gorilla Robot III are shown in Fig. 2. The mechanical structure is designed as follows: 6 DOF for legs, 5 DOF for arms, and 2 DOF for lumbar, and each joint is actuated by AC servomotor. This robot has four force sensors in each end of limb. As a sensor for recognition of environment, a laser range finder (LRF) is mounted perpendicular to the ground on neck of the robot. Its angular resolution is 0.36 [deg], minimum accuracy measurement is 10 [mm], and scan angular range is 240 [deg].

B. *Type of Locomotion*

In biped walking, the dynamics of robot is modeled as a 3D inverted pendulum as shown in Fig. 3(a), following the work in [11]. The supporting point of the pendulum is assumed to be point-contact. Then, only the heeling force and the gravity act on Center of Gravity (COG). To realize

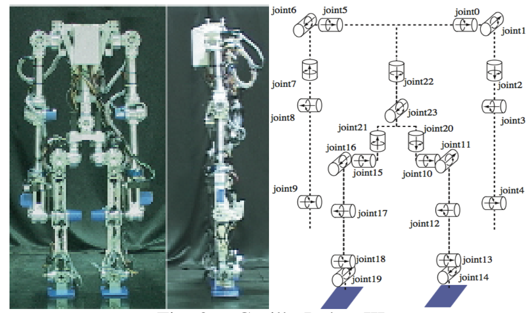
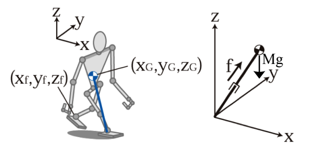
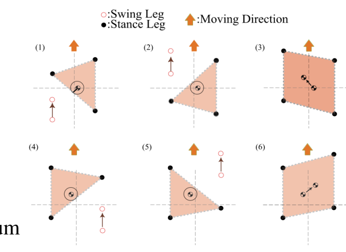


Fig. 2. Gorilla Robot III



(a) Robot modeled as 3D pendulum



(b) Intermittent Crawl Gait

Fig. 3. Type of locomotion

adjustment of moving speed, trajectory of COG is calculated sequentially according to the speed. For the quadruped walking, an intermittent crawl gait is employed [12]. In this mode, the robot moves by the following phase: 1) left rear leg, 2) left front leg, 3) COG, 4) right rear leg, 5) right front leg, 6) COG, (see Fig. 3(b)). Although this gait is stable, fast movement cannot be realized since the robot move the COG and each leg separately. The transition between biped and quadruped walking is made keeping the static balance.

III. EVALUATING FALLING RISK

We utilize "Falling Risk" using Bayesian Network (BN) as a novel indicator of uncertainties and the robot evaluates Falling Risk [7]. Fig. 4 shows the BN model. Nodes $\boldsymbol{X}$ show measured information, $\boldsymbol{Y}$ are direct falling factors, and Z means whether the robot will fall or not. This tree structure is obtained by offline learning via simulations. All nodes of the model take either TRUE $= 1$ or FALSE $= 0$.

Nodes $\boldsymbol{X}$ correspond to the measured information $\boldsymbol{x}$; $x_1 \sim x_4$ are error information that are assumed to be Gaussian noise, and x_5, x_6 are environmental information for which the higher values are expected in the lower probability.

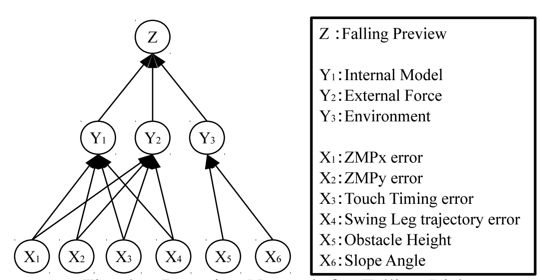


Fig. 4. Bayesian Network for Falling Risk

Hence, $\boldsymbol{x}$ is defined by the following normal distribution.

$$f(x_i) = \frac{1}{\sqrt{2\pi}\sigma_i}\exp\left(\frac{-x_i^2}{2\sigma_i^2}\right). \quad (1)$$

Nodes $\boldsymbol{X}$ are considered as existence probability of $\boldsymbol{x}$. Probability of $X_i = 1$ is given by

$$\Pr(X_i = 1) = \int_{-x_i}^{x_i} f(x)dx = \operatorname{erf}\left(\frac{x_i}{\sqrt{2\sigma_i^2}}\right). \quad (2)$$

The parameters σ_i in eq.(2) is defined by the fact that probability of including x_i in the range $[-2\sigma_i, 2\sigma_i]$ is equal to 95.45%, and the assumption that x_i exists 95.45% within maximum permissible value $\pm x_{\mathrm{imax}}$.

$$\sigma_i := \frac{x_{\mathrm{imax}}}{2}, \quad (3)$$

where x_{imax} is defined according to each gait: $x_{1\mathrm{max}}, x_{2\mathrm{max}}, x_{4\mathrm{max}}$ are determined by the locomotion design, such as the size of the support polygon, and $x_{3\mathrm{max}}, x_{5\mathrm{max}}, x_{6\mathrm{max}}$ are decided by limits of the locomotion, such as the slope angle.

Falling Risk S is defined as probability of $Z = 1$.

$$S(k, a) := \sum_{\boldsymbol{Y}} \Pr(Z = 1|\boldsymbol{Y}, a) \prod_{i=1}^{3} \Pr(Y_i), \quad (4)$$

$$\Pr(Y_i) = \sum_{\boldsymbol{X}_{\boldsymbol{Y_i}}} \Pr(Y_i|\boldsymbol{X}_{\boldsymbol{Y_i}}, a_I) \prod_{\boldsymbol{X}_{\boldsymbol{Y_i}}} \Pr(X), \quad (5)$$

where $\boldsymbol{X}_{\boldsymbol{Y_i}}$ denotes the parent nodes of Y_i, a is one of the available gaits, and a_I is the selected gait at walk-cycle I. Marginal probability of Y_i is assumed to be the same regardless of the selected gait since any locomotion has the same links between $\boldsymbol{Y}$ and $\boldsymbol{X}$, and the similar CPT of $\Pr(Y_i|\boldsymbol{X}_{\boldsymbol{Y_i}}, a)$. According to eq.(4), the robot can evaluate Falling Risk for each gait.

IV. STRATEGY I. LOCOMOTION SELECTION

We have proposed a selection algorithm of gait based on Falling Risk as an indicator of uncertainties and moving speed. As a prerequisite, we define an environmental model for the state of Falling Risk and the gait as Semi-Markov Decision Process (SMDP) [13]. Unlike normal Markov Decision Process (MDP) [14], SMDP has the peculiarity in that timing of decision-making is limited. Actually, the robot can select a gait only during double support phase. In SMDP, evaluation functions are commonly accumulated. Considering such a peculiarity, as an indicator of selection, we define locomotion reward as follows:

$$\begin{aligned} R(k, a) &:= \gamma_R R(k-1, a) + R_0(a) \\ &+ \frac{1-\gamma_S}{1-\gamma_S^{N+1}} \sum_{n=0}^{N} \gamma_S^n V_0(S(k+n), R_0(a)). \end{aligned} \quad (6)$$

Where, each term is given as follows:

$$R_0(a) := \begin{cases} C_R v(a) & (a = a_I) \\ \gamma_T C_R v(a) & (a \neq a_I) \end{cases}, \quad (7)$$

$$V_0(S, R_0(a)) := -\frac{R_0(a)}{1 + \mathrm{e}^{r(1/2-S)}}. \quad (8)$$

Equation (7) is the basic reward that indicates the value for efficiency that a gait has using moving speed $v(a)$, and eq.(8) is the state penalty for Falling Risk S. r is decided as $V_0(S = 0) \simeq 0$. $\gamma_R, \gamma_S, \gamma_T$ are introduced to determine how to select; γ_R is the weight of the past, γ_S is the weight of the future, and γ_T is the degree of influence by transition. N is the maximum future time step for the prediction of Falling Risk. C_R is a coefficient for avoiding rounding errors.

The sum of R_0 and V_0 satisfies $0 < R_0 + \sum V_0 < R_0$. Therefore, the state penalty is dominant if locomotion reward is close to 0 and basic reward is dominant if locomotion reward is close to basic reward. This method guarantees that every gait can be selected since all of the possible range of $R(a)$ always partially overlap by aligning the lower limit.

Using greedy algorithm [15], the robot selects the gait depending on the maximum locomotion reward a^* as follows:

$$a^* := \arg\max_a R(k_{\mathrm{sel}}, a). \quad (9)$$

Where, k_{sel} is the time step when the robot selects a gait. After the selection, the robot initializes all locomotion reward in order to eliminate the effects of past state.

V. STRATEGY II. SPEED ADJUSTMENT

A. Extension of locomotion reward

We extend the locomotion selection algorithm to adjust moving speed for low-level uncertainties that can conform without locomotion selection. Since adjusting moving speed can be done even in one leg-supporting phase, the robot can select the speed periodically. That is, the environment is modeled by MDP not SMDP [13], [14]. It is expected that Falling Risk may change with varying speed in movement. However, we cannot predict the behavior of the Falling Risk unless the node of moving speed is added in BN. Then, we assume that true Falling Risk S^* given by considering the moving speed is represented by a linear approximation.

$$S^* \simeq S + C_v \frac{dv}{v_b}. \quad (10)$$

Where, C_v is the degree of influence on moving speed for the state. v_b is basis speed, dv is difference between the current speed and basis speed. By eq.(10), we can predict Falling Risk against the change of speed with small calculation cost.

Equation (8) is modified as follows:

$$\begin{aligned} V_0(S, R_0) &:= -\frac{R_0}{1 + \exp r\left\{D - \left(S + C_v \frac{dv}{v_b}\right)\right\}}, \quad (11) \\ &= -\frac{R_0}{1 + \exp r\left\{D^* - S\right\}}. \quad (12) \end{aligned}$$

Where, D is the parameter for the demand of stability (in this study, $D = 1/2$). Here, we divided the area of D into the following three cases.

(i) $D < 1/2$: The robot tries to move securely.
(ii) $D = 1/2$: The robot wants to pursue the maximum efficiency without a falling.
(iii) $D > 1/2$: The robot has no problem even if a falling starts since the robot can avoid the falling.

As another interpretation of eq.(11), eq.(12) shows the relationship between the speed and the demand of stability, where $D^* := D - C_v \frac{dv}{v_b}$. If the robot increases the speed, D^* will correspond to the case (i) and the robot aims at secure movement. Also, decreasing the speed updates D^* such that it corresponds to the case (iii), which allows the robot to avoid a falling.

B. Adjustment of moving speed

The robot can predict changes of locomotion reward by adding/subtracting minimal speed δv to/from the current speed v_n under the assumption that Falling Risk S is not fluctuated by changing speed. Using greedy algorithm [15], the robot selects the maximum locomotion reward obtained at that time to update moving speed.

$$d^* := \arg\max_i R(k, a, v_n + i\delta v) \quad (i = -1, 0, 1), \tag{13}$$

$$v_n \Leftarrow v_n + d^* \delta v. \tag{14}$$

As for basis speed v_b, if we keep v_b from the beginning of the movement, adjustment of moving speed will be restricted only in a very narrow range, and a sufficient effect cannot be expected. We considered that improvement of mobility by second strategy was more effective by updating the basis speed in a constant cycle (when the locomotion selection). Therefore, we use Exponential Moving Average (EMA) defined as eq.(15) to update.

$$\mathrm{EMA}(k) = \gamma_v \delta v(k) + (1 - \gamma_v)\mathrm{EMA}(k-1), \tag{15}$$

$$v_b \Leftarrow v_b + \mathrm{EMA}(k_{\mathrm{sel}}). \tag{16}$$

Where, $\gamma_v \in [0, 1]$ is the discount rate from the basis speed.

C. Analysis of adjusting moving speed

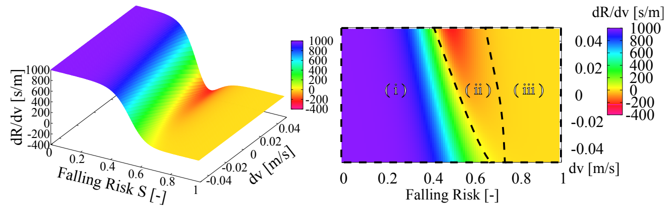
We analyze behavior of adjusting moving speed depending on Falling Risk and the speed. For simplicity, we set the parameters as $\gamma_R = 0, N = 0$. Locomotion reward eq.(6) consists of simply the current component $R_0 + V_0(S(k), R_0)$. Differentiating R with respect to v, we consider that $\frac{dR}{dv}$ is a function for dv and Falling Risk S. However, it is difficult to determine the relationship between S and dv since $\frac{dR}{dv}$ is a complex non-linear equation.

Then, we confirm the qualitative property for $\frac{dR}{dv}$. Fig. 5 shows the graph of $\frac{dR}{dv}$ with respect to Falling Risk and the speed. We divided the graph into the following three regions.

(i) State of Robot is Stable:
The moving speed is expected to increase since $\frac{dR}{dv}$ is positive regardless of the speed. That is why Falling Risk is hardly affected by a change of moving speed.

(ii) State of Robot is Uncertain:
$\frac{dR}{dv}$ tends to be negative particularly when dv is high value. That is, the speed decreases in such cases where current speed is large for the basis speed. Even if the speed is lower than the basis speed, further slow down is expected by relatively high Falling Risk.

(iii) State of Robot is Unstable:
At any moving speed, the robot will maintain the current speed since $\frac{dR}{dv}$ is almost zero and variation of R by δv is zero because of rounding errors. In this case, the robot will stabilize by the locomotion selection. This behavior can be interpreted as that the robot selects the locomotion selection strategy due to exceeding the support level by adjusting the speed.



(a) Overhead view (b) Top view

Fig. 5. Analysis of the behavior of adjusting moving speed

TABLE I

PARAMETERS FOR SIMULATION AND EXPERIMENT

Symbol	Meaning	Value
C_R	Coefficient for avoidance of rounding error	1000
γ_R	Discount rate for impact of past	0.3
γ_S	Discount rate for reliability of estimation	0.7
γ_T	Discount rate for ease of transit	0.8
N	Maximum estimated time step	16
C_v	Coefficient for influence of Falling Risk	v_b
γ_v	Discount rate for basis speed	0.5
T_{trans}	The time for transition	12.0 [s]
$v_{b0}(b)$	Initial moving speed of biped walking	0.080 [m/s]
$v_{b0}(q)$	Initial moving speed of quadruped walking	0.025 [m/s]
v_n	Actual moving speed	$[\frac{1}{2}v_{b0}, \frac{3}{2}v_{b0}]$

From the above analysis, speed adjustment by extended locomotion selection algorithm deals with uncertainties at the level when it is difficult to determine whether executing locomotion selection. In addition, the robot speeds up as much as possible when uncertainties are small. It is suggested that second strategy does not react to the level of uncertainties that is dealt with locomotion selection. The robot can execute locomotion selection as soon as possible, avoiding a slow down of moving speed against large Falling Risk.

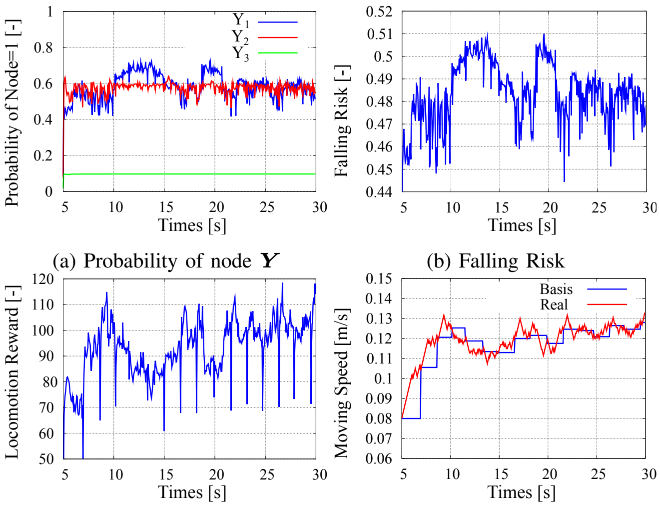
D. Simulation

1) Simulation Conditions: We use a dynamics simulator OpenHRP3. Table I shows the parameters of two strategies that are moving speed, transition time, and range of moving speed. C_v is set as the basis speed v_b, so eq.(10) becomes simply $S + dv$. In this simulation, we aim to confirm the behavior fits the analysis result; the robot increases the speed while the stable state or decreases the speed while the uncertain state. Two bumps are set on walking space as small uncertainties: height 10 [mm] and 500 [mm] diameter.

2) Simulation Result: Fig. 6 and Fig. 7 show the simulation results. The robot started walking from 5 [s].

From 10 [s] to 15 [s] and from 17 [s] to 20 [s], the robot walked on a rough terrain. When the robot was on the rough terrain, $\Pr(Y_1 = 1)$ indicating the internal model error increased, and accordingly Falling Risk became higher. As a result, the robot decreased the speed in order to reduce the inertia and improve tracking the COG trajectory, since the state of the robot was in the region (ii) of Fig. 5.

On the other hand, the robot walked on flat during other times. When the robot was on flat even if small flat like



(a) Probability of node $\boldsymbol{Y}$

(b) Falling Risk

(c) Locomotion reward

(d) Real and basic moving speed

Fig. 6. Simulation results

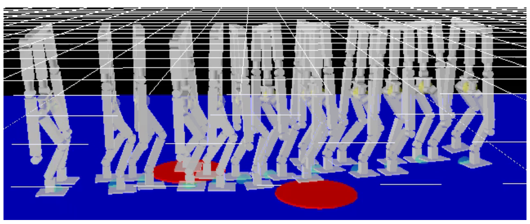


Fig. 7. Snapshots of simulation

from 15 [s] to 17 [s], Falling Risk was relatively small and did not reach allowable Falling Risk. As a result, the robot prioritized moving efficiency by speed up, since the state of the robot was in the region (i) of Fig. 5.

From these two types of behavior, we considered that the simulation results matched the analysis results. That is, the robot obtained the maximum moving efficiency and stability in the range locomotion had by selecting moving speed so as to maximize locomotion reward at that time.

VI. EXPERIMENT USING PHYSICAL HARDWARE

A. Experimental Condition

We verify that the two proposed strategies can work on a degree of uncertainties and improve total mobility. Available gaits are biped / quadruped walking and we use the Gorilla Robot III as described in Section II. Experimental parameters are the same as simulation shown in Table I.

The robot scans the ground surface with LRF before walking as shown in Fig. 8. Here, target distance is 2.5 [m]. Cushions are paved and two steps are placed on the experimental course in order to make partial rough terrain. Cushions ($\sim$ 10 [mm] height) cannot be scanned since the resolution of LRF is 10 [mm]. We can expect that the robot removes by speed adjustment since uncertainties by rough terrain are small. If the robot did not use speed adjustment, we confirmed falling. The presence of two steps on the path was recognized by the scanned data from the LRF (about 50 [mm] height). The robot is hard to go through the two steps in

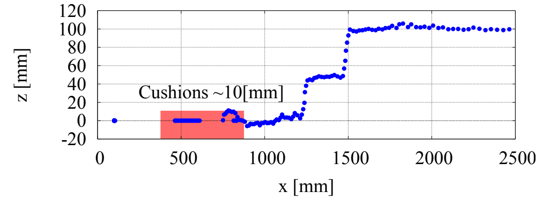


Fig. 8. Scanned ground surface by LRF

biped walking; thus the transition from biped to quadrupled walking will be activated to cope with the situation.

B. Experimental Results

Fig. 9 shows the snapshots of the experiment. Fig. 10 shows experimental data: Probabilities of $\boldsymbol{Y} = 1$, Falling Risk $S(a)$, locomotion reward $R(a)$, and moving speed v_n, v_b. The plots of $R(a)$ vs. t are shown the locomotion reward when the robot selects gait. If the maximum locomotion reward changes, the robot changes gait. During transition motion, the robot does not estimate these parameters.

1) Before steps: The robot started walking at 5 [s]; then the robot immediately increased the walking speed since the initial basis speed was low and the state was stable enough. After 8 [s], the robot stepped into the rough terrain area. Change in Falling Risk was hardly seen, although probability of Internal model factor Y_1 increased. That is the result that the robot increased the stability by suppressing the increase of moving speed by second strategy. The robot walked over the rough terrain in 13 [s], and increased walking speed slightly. However, the robot kept moving speed based on rising of Falling Risk and decreasing of locomotion reward since two steps are at hand. The robot selected quadruped walking in order not to fall in front of the two steps.

2) While steps: Moving speed rose to the limit shown in Table I since two steps are small uncertainties for quadruped walking. At about 48 [s], two steps finished as environmental factor Y_3, and actually the other factors Y_1, Y_2 were not large, the robot selected biped walking in favor efficiency.

3) After steps: After 68 [s], we could confirm that the robot increased walking speed on flat terrain because of small uncertainties. However, stagger of the robot was occurred since moving speed was too high in 74 [s]; and then probability of $Y_1 = 1$ was increased. Accompanying it, the robot decreased walking speed in order to stabilize and kept move without falling.

4) Efficiency comparison: In previous work [7], the robot can use only first strategy and will select quadruped walking when the robot enters the rough terrain, and traveling time will take about 110 [s] (the case A). In this experimental result, traveling time took 80.6 [s] with two strategies (the case B). In comparing the case A and B, second strategy improved moving speed about 37.5%.

These results verified that the robot achieved the high mobility using two strategies; each strategy enhances adaptability against the changes of uncertainties suitable for it.

Fig. 9. Snapshots of experiment; cushions during 8 [s] $\sim$ 13 [s], two steps during 35 [s] $\sim$ 48 [s], transition in 68 [s], stumbling in 74 [s] but recovered

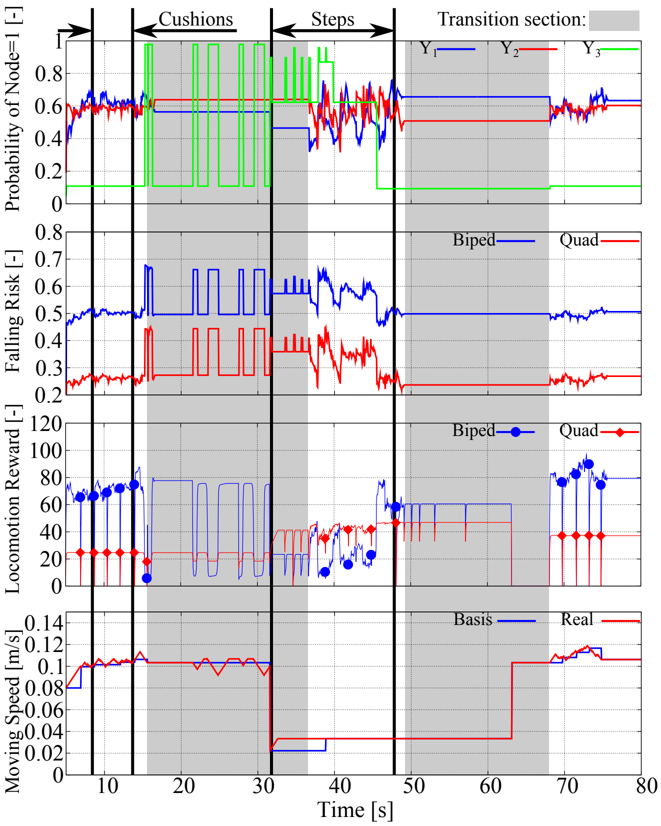


Fig. 10. Experimental results: $\Pr(\boldsymbol{Y} = 1)$, Falling Risk $S(a)$, locomotion reward $R(a)$, and moving speed v_n, v_b. The plots of $R(a)$ vs. t are $R(k_{\mathrm{sel}}, a)$. During transition section, these parameters are not estimated.

VII. CONCLUSIONS AND FUTURE WORKS

We proposed two strategies for improvement of walking mobility using the selection algorithm. First strategy is locomotion selection; the robot selects gait that has the maximum locomotion reward. Second strategy is adjustment of moving speed based on extended locomotion selection algorithm. We aimed to determine the optimal speed consciously and autonomously, extending our proposed algorithm to selection of moving speed. This extension is realized by the assumption that the state penalty depends on not only Falling Risk but also the moving speed. We verified that the behavior of second strategy is as aimed by the analysis and simulation for second strategy. We confirmed by the physical experiment that two strategies removed uncertainties suitable for each strategy level without competing, and improved moving speed about 37.5% from the case of only using gait selection strategy. Furthermore, This realization by the extension of proposed locomotion selection algorithm verified not only improving mobility, but also confirmation of the usefulness and extensibility of proposed method.

Thus, we will generalize proposed selection algorithm such that that algorithm can be apply to other decision-making problem. In addition, the objectivity of evaluating Falling Risk is very important factor in proposed selection algorithm. We will make the robot learn Conditional Probability Table of Bayesian Network in real time.

REFERENCES

[1] T. Fukuda, Y. Hasegawa, K. Sekiyama, and T. Aoyama, *Multi-locomotion Robotic Systems: New Concepts of Bio-inspired Robotics*. Springer-Verlag, 2012.

[2] F. Kanehiro, T. Yoshimi, S. Kajita, M. Morisawa, K. Fujiwara, K. Harada, K. Kaneko, H. Hirukawa, and F. Tomita, "Whole body locomotion planning of humanoid robots based on a 3d grid map," in *Proceedings of the IEEE International Conference on Robotics and Automation*, 2005, pp. 1084–1090.

[3] K. Ogata, K. Terada, and Y. Kuniyoshi, "Real-time selection and generation of fall damage reduction actions for humanoid robots," in *IEEE-RAS International Conference on Humanoid Robots*, 2008, pp. 233–238.

[4] K. Toda, M. Sagara, and K. Tomiyama, "Sensor-based biped gait generation scheme for humanoid - implementation and evaluation," in *IEEE-RAS International Conference on Humanoid Robots*, 2004, pp. 656–671.

[5] R. Renner and S. Behnke, "Instability detection and fall avoidance for a humanoid using attitude sensors and reflexes," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2006, pp. 2967–2973.

[6] T. Aoyama, T. Kobayashi, Z. Lu, K. Sekiyama, Y. Hasegawa, and T. Fukuda, *Injury and Skeletal Biomechanics*. InTech, 2012, ch. Locomotion Transition Scheme of Multi-Locomotion Robot, pp. 21–36.

[7] T. Kobayashi, T. Aoyama, K. Sekiyama, Z. Lu, Y. Hasegawa, and T. Fukuda, "Locomotion selection of multi-locomotion robot based on falling risk and moving efficiency," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2012, pp. 2869–2874.

[8] H. Kimura, Y. Fukuoka, and A. H. Cohen, "Fast biped walking with a sensor-driven neuronal controller and real-time online learning," *The International Journal of Robotics Research*, vol. 25, no. 3, pp. 243–259, 2006.

[9] T. Geng, B. Porr, and F. Wörgötter, "Adaptive dynamic walking of a quadruped robot on natural ground based on biological concepts," *The International Journal of Robotics Research*, vol. 26, no. 5, pp. 475–490, 2007.

[10] S. Inagaki, H. Yuasa, T. Suzuki, and T. Arai, "Wave cpg model for autonomous decentralized multi-legged robot: Gait generation and walking speed control," *Robotics and Autonomous Systems*, vol. 54, no. 2, pp. 118–126, 2006.

[11] S. Kajita, F. Kanehiro, K. Kaneko, K. Fujiwara, K. Yokoi, and H. Hirukawa, "Biped walking pattern generation by a simple three-dimensional inverted pendulum model," *Advanced Robotics*, vol. 17, no. 2, pp. 131–147, 2003.

[12] H. Tsukagoshi, S. Hirose, and K. Yoneda, "Maneuvering operations of a quadruped walking robot on a slope," *Advanced Robotics*, vol. 11, no. 4, pp. 359–376, 1997.

[13] S. J. Bradtke and M. O. Duff, "Reinforcement learning methods for continuous-time markov decision problems," in *Advances in Neural Information Processing Systems*, 1994, pp. 393–400.

[14] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, Inc., 1994.

[15] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*. MIT press, 1990.