

Motor Control Optimization of Compliant One-Legged Locomotion in Rough Terrain

Fumiya Iida and Russ Tedrake

Abstract—While underactuated robotic systems are capable of energy efficient and rapid dynamic behavior, we still do not fully understand how body dynamics can be actively used for adaptive behavior in complex unstructured environment. In particular, we can expect that the robotic systems could achieve high maneuverability by flexibly storing and releasing energy through the motor control of the physical interaction between the body and the environment. This paper presents a minimalistic optimization strategy of motor control policy for underactuated legged robotic systems. Based on a reinforcement learning algorithm, we propose an optimization scheme, with which the robot can exploit passive elasticity for hopping forward while maintaining the stability of locomotion process in the environment with a series of large changes of ground surface. We show a case study of a simple one-legged robot which consists of a servomotor and a passive elastic joint. The dynamics and learning performance of the robot model are tested in simulation, and then transferred the results to the real-world robot.

I. INTRODUCTION

There has been an increasing interest in the use of passive dynamics to achieve energy efficient and rapid movement of robotic systems. It was shown, for example, that bipedal walking can be very efficient when robots exploit physical body dynamics [1]. While most of the underactuated legged robots are functional only in relatively simple environment (e.g. a flat terrain or ones with small steps), it is a highly challenging problem to control them in rough environment.

Previously, there have been a few approaches proposed for adaptive control architectures of underactuated legged robots that work in rough terrains. The most successful legged robot model is the series of robots called RHex [2], [3]. The unique morphological design of elastic rimless wheels makes these robots capable of maintaining the locomotion process in many extreme ground conditions without any strict and complex control architectures. More careful foot placement of legged robots was also investigated in the past. The pioneering work of one-leg hopping robot demonstrated locomotion over a large obstacle by using a simple feedback control [4]. A series of investigations on the adaptive motor controller (e.g. a simple sensory feedback exploiting dynamics [5], [6], the studies of central pattern generator models [7], [8], [9], [10], and a reinforcement learning approach [11]) also demonstrated adaptive responses in uneven terrains. These approaches, however, still suffer from the stability in complex environment, and they are

F. Iida and R. Tedrake is with Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, 32 Vassar Street, Cambridge, 02139 MA, USA, iida@csail.mit.edu, russt@mit.edu

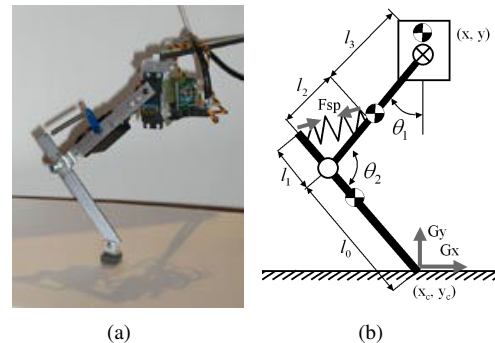


Fig. 1. (a) A photograph of the robot platform and (b) a schematic of the simulation model. The open circle denotes a passive joint and the circle with a cross inside denotes a joint controlled by the servomotor.

still not able to deal with a series of large changes on the ground. More specifically, these approaches require some stabilization periods in the locomotion process after dealing with a large change. In general, the body dynamics was so far exploited only to stabilize the locomotion process, and we still do not fully understand how to configure the dynamics such that it can achieve a series of good foot placement in complex environment.

In general, it is a highly challenging problem to achieve a series of good foot placement because every leg step in underactuated legged robots is dependent on the other preceding leg steps. The goal of this paper is to propose a simple optimization strategy of motor control in underactuated legged robots. This approach considers how to optimize motor control parameters in order to achieve reasonable foot placement in complex environment with a series of large changes of the ground condition. An underactuated one-legged robot is used in simulation and in the real world in order to evaluate the performance of the proposed approach.

In the next sections, we first introduce a model of legged robot, and the dynamics of this model will be analyzed. The learning architecture and its performance evaluation will then be shown, together with a real-world robot experiment.

II. ONE-LEG ROBOT MODEL

Complex anatomical structure is a fertile basis of animals' adaptive behavior. Likewise, "good" mechanical structures and material properties are an important prerequisite for adaptive behavior of autonomous robotic systems. In this section, we explain the mechanical design of a single-legged robot that exploits passive dynamics and elasticity for forward hopping locomotion. The model of a single-

TABLE I
SPECIFICATIONS OF THE ROBOT'S MECHANICAL DESIGN

Parameter	Description	Value
l_0	length of lower leg segment	150 mm
l_1	spring attachment	30 mm
l_2	spring attachment	30 mm
l_3	length of upper leg segment	120 mm
l_4	length of body	110 mm
M	mass of the robot	0.80 kg
θ_{2rest}	rest angle of the leg segments	120 deg
K_{sp}	spring constant	60 kg/m

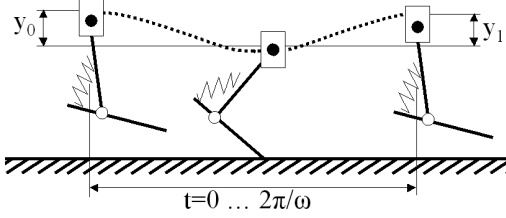


Fig. 2. Analysis of dynamics during one cycle of motor oscillation.

legged robot investigated in this paper is inspired from the previous biomechanics studies [12], [13], [14], and it has been previously shown that it requires very simple motor control for rapid locomotion behavior [15]. This control scheme can also be extended for the other types of locomotion including walking, hopping, and running in two- and four-legged locomotion [6], [15], [16], [17].

Figure 1 shows one of the simplest legged robot models. This single-leg robot consists of one position-controlled motor at the hip joint and two limb segments connected through an elastic passive joint. The behavior of this robot can be described as:

$$M\ddot{\mathbf{q}} = \mathbf{V}(\dot{\mathbf{q}}) + \mathbf{E}(\mathbf{q}, \dot{\mathbf{q}}) + \mathbf{G} + \mathbf{T} \quad (1)$$

where $\mathbf{q} = [x \ y \ \theta_1 \ \theta_2]^T$. \mathbf{V} , \mathbf{E} and \mathbf{G} are velocity and elasticity dependent forces, and gravitational force, respectively. \mathbf{T} is the motor torque provided by the motor control explained below. The velocity and elasticity dependent forces include damping and elasticity elements in the spring and ground reaction force. In the simulation experiments shown in the next sections, the force generated in these tension springs are calculated as:

$$F_{sp} = K_{sp}(d_2 - d_{2rest}) - D_{sp}\dot{d}_2 \quad (2)$$

where K_{sp} and D_{sp} are the spring and damping constants, and d_2 and d_{2rest} represent the length and the rest length of the spring that can be determined by θ_2 and θ_{2rest} .

The ground reaction forces are calculated based on a spring-damper ground interaction model studied in biomechanics [18].

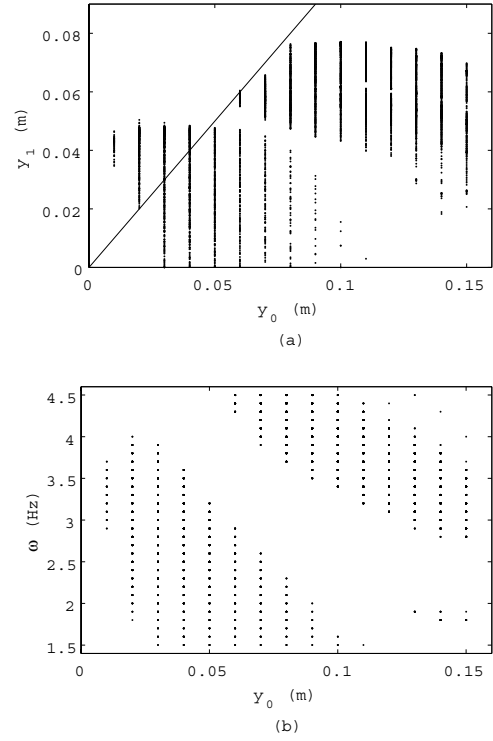


Fig. 3. Single step dynamics of the simulation model. (a) the hopping heights at the second apex (y_1) are plotted with respect to the initial heights (y_0), and (b) shows the stable range of hip joint oscillation frequency.

$$G_y = a|y_c|^3(1 - b\dot{y}_c) \quad (3)$$

$$G_x = \begin{cases} \mu_{slide}G_y \frac{\dot{x}_c}{|\dot{x}_c|} & \mu_{slide}G_y \frac{\dot{x}_c}{|\dot{x}_c|} > \mu_{stick}G_y \\ F_{xc} & \mu_{slide}G_y \frac{\dot{x}_c}{|\dot{x}_c|} \leq \mu_{stick}G_y \end{cases} \quad (4)$$

where x_c and y_c denote horizontal and vertical distances of the foot contact point on the ground surface, respectively. F_{xc} represents the force required to prevent the contact point sliding. There is only one ground contact point defined at an end of the lower segment in this model.

This system requires only a simple motor oscillation with no sensory feedback to stabilize itself into a periodic hopping behavior. Therefore the \mathbf{T} term in equation (1) is calculated to achieve PD control of the following sinusoidal oscillation.

$$\theta_1(t) = A \sin(2\pi\omega t) + B \quad (5)$$

In the following simulation and robot experiments, we consider the control parameter of frequency ω only, and the amplitude and offset parameters are fixed at $A = 35$ and $B = 40$ degrees.

For the sake of real-world implementation and experiments, the dimension of this model is scaled down as shown in Table I. And to facilitate the analysis, this model is restricted to the motion within a plane, and no rotational movement (roll or pitch) of the body segment is considered. In the following simulation experiments, we implemented

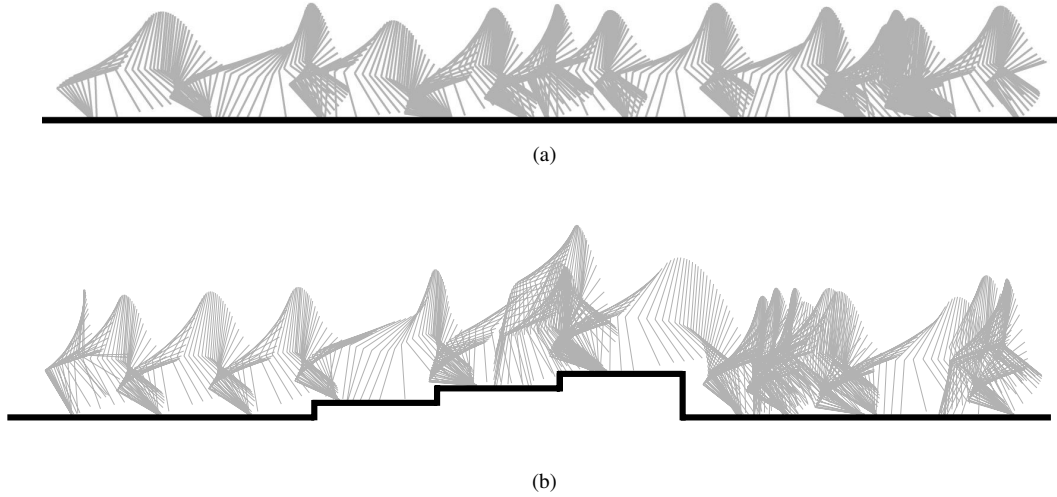


Fig. 4. Simulated behavior of the robot (a) in a flat terrain and (b) in a rough terrain. The sequence of motor frequencies are optimized by the proposed algorithm explained in the text.

this model to MathWorks Matlab together with the SimMechanics toolbox.

III. ANALYSIS OF BODY DYNAMICS

Although this robot model is simple, it exhibits relatively complex behavior patterns through the compliant-viscous interactions of the passive joint and the ground surface. In this section, we explain the body dynamics of the robot model to understand the basic characteristics and capabilities of the system. This analysis will be used to determine the parameter space for the optimization process later.

The following analysis was conducted in simulation, and we focus on the behavior during one leg step of this robot model as shown in Figure 2. Although this model has eight state variables (i.e. \mathbf{q} and $\dot{\mathbf{q}}$), we approximate the basic dynamics by examining three state variables $[y \ \dot{x} \ \theta_1]$. More specifically, by considering the dynamic interaction from one apex of flight phase to the subsequent apex, we assume that x can be arbitrary, $\theta_2 = \theta_{2rest}$ and $\dot{\theta}_2 = 0$ as the passive joint stays at the rest length of the spring. Therefore, in this experiment, we conducted the simulation of one motor oscillation cycle with various initial heights y_0 , horizontal velocity \dot{x} , and motor control frequency ω .

Figure 3 shows the single step dynamics of the simulation model, and the plots indicate the initial condition of hopping height with which the system successfully reached to the next apex height. During this simulation, the behavior patterns resulting in more than one contact points touching the ground surface, and/or the passive joint exceeding large threshold angles were regarded as an unstable locomotion process, thus they were ignored. As shown in Figure 3(a), this model is able to achieve stable bounding with a relatively broad range of initial heights y_0 . Namely, it is able to maintain the hopping gait when the apex height is between 0.02 and 0.06 m in particular, because the plots are around the solid line in the figure. Even with the high initial height, the subsequent

apexes result in lower heights which makes the system capable of maintaining locomotion processes eventually.

Note that there are two large clusters of plots in Figure 3(a) and (c), which indicates two different kinds of gait patterns. The cluster at the smaller initial heights and frequency corresponds to the gait, in which the foot touches down the ground while the leg is swinging back. The other cluster at the larger initial heights and frequency corresponds to the gait pattern of the foot touchdown while the leg is swinging forward. For the sake of nomenclature, we call the former pattern “Gait 1” and the latter one “Gait 2”.

From this analysis, we can conclude that the proposed system is theoretically capable of maintain the locomotion process with any values of state variables examined in this experiment. Considering the real-world implementation, however, we decided to employ the motor frequency parameter between 1.5 and 3.5 Hz, because Gait 2 generally induces undesirable large impact force of touch down at the joints.

IV. OPTIMIZATION OF CONTROL POLICY

As shown in the previous section, the dynamics of the proposed model can be used for maintaining locomotion process in relatively broad range of state variables. It is, however, necessary to adjust the motor control parameter to deal with rough terrains. This section explains an optimization algorithm and its performance.

A. Optimization Algorithm

In order to minimize the dimension of controller optimization, the following scheme considers only one control parameter, i.e. motor frequency ω . Moreover, we consider the optimization process can explore the frequency parameter at every leg step, and it is not allowed to change it during the cycle. The target control policy of the learning process, therefore, is $\pi = \pi(\omega_1, \omega_2, \dots, \omega_n)$.

In the learning process, we utilized a stochastic search method by using a two-dimensional probability matrix

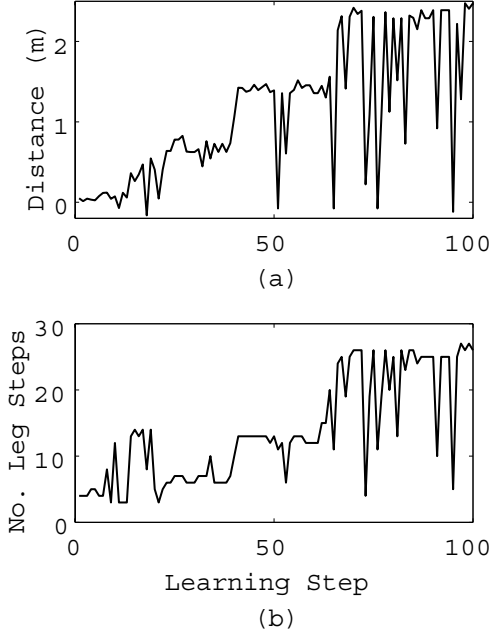


Fig. 5. A typical optimization process of the proposed algorithm during running on a flat terrain (in simulation). (a) An increasing travelling distance during the learning process and (b) the number of leg steps.

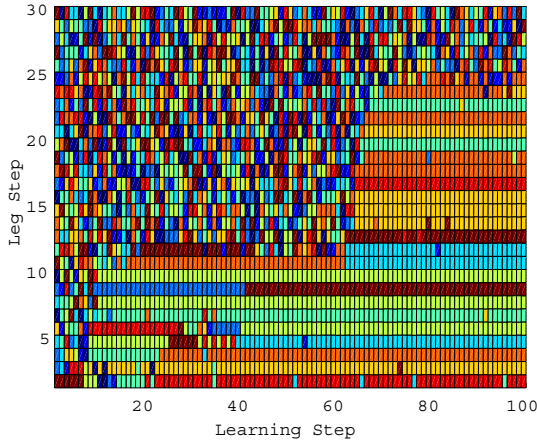


Fig. 6. A learning process of motor control policy. The color in each tile indicates the oscillation frequency of motor at the leg step N.

$Q(N, a)$ with which the learning process determines a oscillation frequency ω_N at leg step N . Based on the dynamics analysis in the previous section, we defined a discrete set of 10 actions as $\mathbf{a}_{i=1, \dots, 10} = [1.5, 1.7, \dots, 3.3]$ Hz. Note that, unlike a standard reinforcement learning, we do not use sensory states as a reference of this probability matrix, but the number of leg step cycle N is used instead.

At the beginning of learning episode t , the learning process determines a control policy π^t based on the Q matrix by using a roulette rule function F_{roulette} .

$$\pi^t = F_{\text{roulette}}(Q^t(i, a), \epsilon) \quad (6)$$

$$i = 1, \dots, N$$

where the parameter ϵ determines the degree of stochastic search: with a higher ϵ value, for example, the process selects more random actions based on the Q values.

After running an episode with the control policy π^t , a reward value of the leg step N is evaluated as follows:

$$R^t(i) = \begin{cases} -5.0 & : i = \text{FailedStep} \\ \text{FinalDistance} & : i \neq \text{FailedStep} \end{cases} \quad (7)$$

In the reward evaluation, the action at the leg step N which results in a failure locomotion process receives a negative reward. Namely, the probability to select ω_N in the next episode will be decreased.

These rewards are then used to update the Q matrix as follows:

$$Q^{t+1}(i, a_i) = (1.0 - \alpha) \cdot Q^t(i, a_i) + \alpha \cdot (R^t(i) + \gamma \cdot \text{mean}(Q^t(i+1, a))) \quad (8)$$

where α and γ represent learning and discount rates, respectively. Note that the discount term in this equation also evaluates the preceding leg steps from the leg step i . Because the instability of locomotion process often caused by multiple inappropriate leg steps earlier than the actual failed one, this discount term improves the learning performance significantly.

B. Simulation Result in Flat Terrain

We first explain a series of simulation experiments of the learning algorithm in a flat terrain. We conducted 100 episodes of optimization processes, and repeated this process with 100 different random seeds.

Figure 4 illustrates the kinematic trajectories of body segments in simulation experiments. A locomotion behavior with an optimized control policy is shown in Figure 4(a), where the optimization process mostly find control policies to maintain forward hopping behavior in a flat terrain. It is interesting, however, that the optimized control policies are generally a set of unsteady values of motor frequencies even in the locomotion on a flat terrain. Namely, the system changes the motor frequency almost every step, which results in unsteady behavior patterns. As shown in Figure 4(a), for example, it slows down sometimes after a few forward movement. The optimized control policy showed, however, a better performance (0.48 m/sec) than a steady control policy running at a fixed motor oscillation frequency 0.27 Hz (0.18 m/sec).

Figure 5 shows a detailed process of optimization. In this example, the learning process was able to optimize the locomotion behavior for 26 leg steps after 100 episodes, with which the system could travel approximately 2.5 m. It is important to note that this learning process optimizes both maximum travelling distance and locomotion stability:

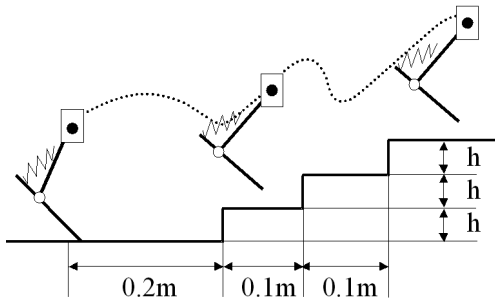


Fig. 7. The environmental conditions with steps for the simulation analysis.

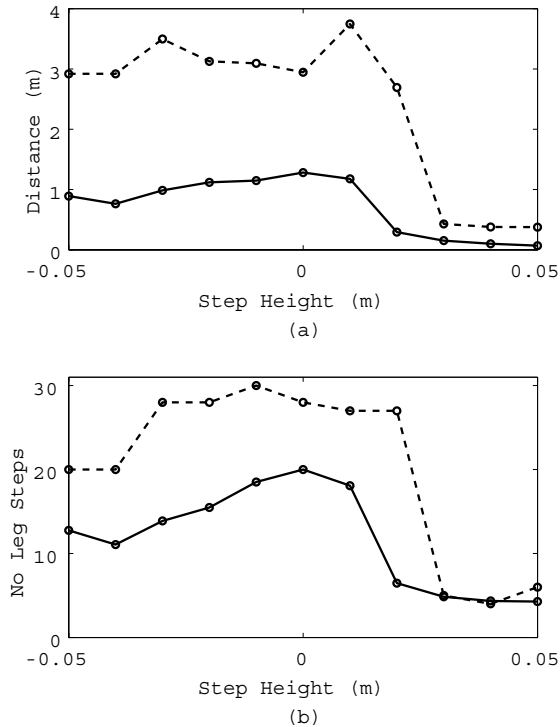


Fig. 8. Optimization of locomotion in different heights of steps. (a) Maximum travelling distance (dashed lines) and mean distance (solid lines) over 50 optimization processes, and (b) the maximum and mean leg steps.

by comparing Figure 5(a) with (b), the number of leg steps increasing while the travelling distance does not (for example, around the learning steps 15 - 20).

The exploration of control policy can be more clearly shown in Figure 6. At the learning steps 15 - 20, for example, the process is optimizing the leg step 11 and 12. The stochastic optimization process optimized for about 50 learning steps, it found a solution to maintain the locomotion process further.

C. Simulation Result in Rough Terrain

The same optimization process can also be applied in rough terrains. One of the optimized results is shown in Figure 4(b), in which the simulated robot successfully climbed up

three steps and jumping down a large gap while maintaining the stable locomotion process. This figure illustrates how the system deals with the changes of environment with multiple leg steps: after climbing up the first and second steps, the simulated robot exhibits multiple leg steps to maintain the stability by hardly moving forward.

The optimization performance in rough terrains was also evaluated in the simulation experiments. As shown in Figure 7, we have conducted the optimization processes in the environment of three steps with various heights. This figure shows the maximum distance and number of leg steps after the simulation of 50 episodes of 100 learning steps. It clearly shows that the system is able to deal with relatively high decreasing steps, although it is not able to cope with the increasing steps more than 0.03 m high.

V. ROBOT EXPERIMENT

The proposed optimization scheme was then tested in the real-world robot. The optimization process was first conducted in simulation and transferred to the controller of the real-world robot. To minimize the difference between the simulation model and the real world environment, we set up the environment with a few rubber sheets as shown in Figure 9. The ground surface contains three steps (20 mm, 25 mm, and 27 mm high, respectively).

Although there are some differences between simulation and the real-world environment especially in the ground interactions, the resultant behavior of the robot is more or less comparable with that of simulation as long as it is a few leg steps. The robot exhibits two leg steps after the first step (i.e. the pictures 4-8 in Figure 9), and then stepping two more steps.

VI. DISCUSSION AND CONCLUSION

This paper presented a motor control optimization scheme of underactuated one-legged robot system. By using an optimization method similar to reinforcement learning algorithm, the proposed approach optimizes a sequence of motor commands of dynamic locomotion, which exploits the body dynamics to deal with a series of large steps on the ground. The method was evaluated both in simulation and the real-world robot and it was demonstrated that the locomotion process can be optimized in relatively short time steps. There are, however, a few points that need to be discussed further.

Although the typical reinforcement learning generally utilizes a complete set of state information in the optimization process which guarantees optimal control policies, the proposed method does not consider any sensory information except for the number of leg steps. Because of the absence of sensory feedback (i.e. the feedback of body state variables and the information about the rough terrain), the learning system does not need to explore the vast state space of the system. While the advantages of this approach lies in the fact that the system is capable of finding dynamic behavioral patterns in relatively short period of learning process, a drawback of this approach is that it requires learning processes whenever the system encounters different

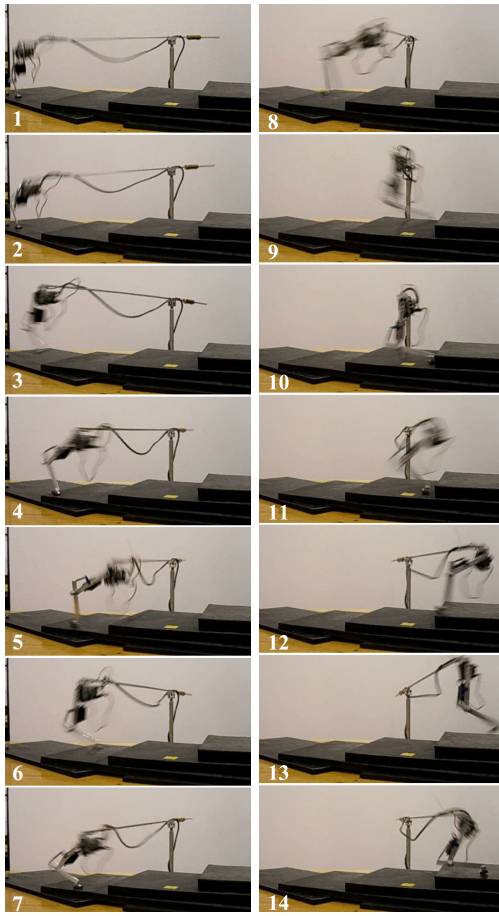


Fig. 9. Behavior of the robot over three steps. The intervals between the pictures are approximately 30 ms.

initial conditions and rough terrains. In order to account for this problem, we are currently working on extending the proposed learning scheme such that the control policy could take the sensory information into account, based on the motor control optimization in this paper.

Alternatively, there are a few additional related work that could be combined with the proposed approach toward more plausible real-world implementation. For example, we still do not know how to implement the optimization process in the real-world robot. In addition, the robot model used in this paper is very simple and it still has many constraints (e.g. the pitch, roll and yaw rotations are fixed), which also needs to be considered further in the future. We expect that a similar optimization scheme can be used for the other related robot models such as biped and quadruped robots [16], [17], [19].

APPENDIX

In the simulation experiments presented in this paper, we used the following parameter values: $K_{sp} = 600(N/m)$, $D_{sp} = 10(N \cdot s/m)$, $a = -2.5e^5$, $b = 3.3$, $\mu_{slide} = 0.5$, and $\mu_{stick} = 0.6$. The gain parameters of the PD controller are: $K_p = 15$ and $K_d = 0.01$.

ACKNOWLEDGMENTS

This work is supported by the Swiss National Science Foundation Fellowship for Prospective Researchers (Grant No. PBZH2-114461).

REFERENCES

- [1] Collins, S., Ruina, A., Tedrake, R., and Wisse, M. (2005). Efficient bipedal robots based on passive dynamic walkers, *Science Magazine*, Vol. 307, 1082-1085.
- [2] Altendorfer, R., Moore, N., Komsuoglu, H., Buehler, M., Brown, H. B., McMordie, D., Saranli, U., Full, R., and Koditschek, D. (2001). Rhex: A biologically inspired hexapod runner, *Autonomous Robots*, vol. 11, 207.
- [3] Saranli, U., Buehler, M., Koditschek, D.E. (2001). RHex: A simple and highly mobile hexapod robot, *The International Journal of Robotics Research*, 20, 616-631.
- [4] Hodgins, J., Koechling, J., Raibert, M. H. (1985). Running experiments with a planar biped. *Third International Symposium on Robotics Research*, Cambridge: MIT Press.
- [5] Takuma, T. and Hosoda, K. (2006). Controlling the walking period of a pneumatic muscle walker, *The International Journal of Robotics Research*, Vol. 25, No. 9, 861-866.
- [6] Zhang, Z.G., Kimura, H. and Fukukoka, Y. (2006). Autonomously generating efficient running of a quadruped robot using delayed feedback control, *ADVANCED ROBOTICS*, Vol.20, No.6, pp.607-629.
- [7] Fukuoka, Y., Kimura, H., and Cohen, A. H. (2003). Adaptive dynamic walking of a quadruped robot on irregular terrain based on biological concepts, *The International Journal of Robotics Research*, Vol. 22, Issue 3, 187-202.
- [8] Taga, G., Yamaguchi, Y., and Shimizu, H. (1991). Self-organized control of bipedal locomotion by neural oscillators in unpredictable environment. *Biological Cybernetics* 65, 147-159.
- [9] Geng, T., Porr, B., Wörgötter, F. (2006). A reflexive neural network for dynamic biped walking control, *Neural Computation*, Vol. 18, No. 5, 1156-1196.
- [10] Matsubara, T., Morimoto, J., Nakanishi, J., Sato, M., Doya, K. (2005). Learning CPG-based biped locomotion with a policy gradient method, *Proceedings of 2005 5th IEEE-RAS International Conference on Humanoid Robots*, 208-213.
- [11] Ueno, T., Nakamura, Y., Takuma, T., Shibata, T., Hosoda, K., and Ishii, S. (2006). Fast and stable learning of quasi-passive dynamic walking by an unstable biped robot based on off-policy natural actor-critic. *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2006)*, 5226-5231.
- [12] Blickhan, R. (1989). The spring-mass model for running and hopping, *J Biomechanics* 22:1217-1227.
- [13] McMahon, T. A., Cheng, G. C. (1990). The Mechanics of Running: How Does Stiffness Couple with Speed?, *J. Biomechanics*, Vol. 23, Suppl. 1, 65-78.
- [14] Seyfarth, A., Geyer, H., Guenther, M., Blickhan, R. (2002). A movement criterion for running, *Journal of Biomechanics*, 35, 649-655.
- [15] Rummel, J., Iida, F., and Seyfarth, A. (2006). One-legged locomotion with a compliant passive joint, *Intelligent Autonomous Systems ? 9*, Arai, T. et al. (Eds.), IOS Press, 566-573.
- [16] Iida, F., Gomez, G. J., and Pfeifer, R. (2005). Exploiting body dynamics for controlling a running quadruped robot. *ICAR 2005*, July 18th-20th, Seattle, U.S.A., 229-235.
- [17] Iida, F., Rummel, J., and Seyfarth, A. (2007). Bipedal walking and running with compliant legs, *IEEE International Conference on Robotics and Automation (ICRA'07)*, 3970-3975.
- [18] Gerritsen, K. G. M., van den Bogert, A. J., and Nigg, B. M. (1995). Direct dynamics simulation of the impact phase in heel-toe running, *Journal of Biomechanics*, Vol. 28, No. 6, 661-668.
- [19] Buchli, J., Iida, F., and Ijspeert, A. J. (2006). Finding resonance: Adaptive frequency oscillators for dynamic legged locomotion, *Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 06)*, 3903-3909.