SICE Annual Conference 2010
August 18-21, 2010, The Grand Hotel, Taipei, Taiwan

Evolution strategies for biped locomotion learning using nonlinear oscillators

Takeshi Uchitane[1], Toshiharu Hatanaka[1] and Katsuji Uosaki[2]
[1]Depaertment of Information and Physical Sciences, Osaka University
Suita, 565-0871, Japan,
[2]Department of Management Information Science, Fukui University of Technology
Fukui, 910-8505, Japan
email: {t-uchitane, hatanaka}@ist.osaka-u.ac.jp, uosaki@fukui-ut.ac.jp

***Abstract*—This paper addresses a tuning method for the parameters in a gait pattern, control systems and nonlinear oscillators of humanoid robots. The phase oscillators with PD controller are used to generate a rhythmic walking pattern, thus a walking pattern is described by many parameters. Using an evolutionary computation, our approach uses only performance evaluation values to tune these parameters. In this paper, we propose a novel evolution strategy that uses mask on the portion of individual to avoid mutation. Numerical simulation studies are carried out to evaluate the performance of the proposed approach by using the RoboCup 3D Soccer Simulator.**

***Keywords*—Central pattern generator, Evolution Strategy, biped locomotion, RoboCup 3D Soccer Simulator**

I. INTRODUCTION

The humanoid robots are receiving much attention to develop robotic systems coexisting with human beings such as livelihood support robots and the rescue robots. And another important issue is to understand human beings by constructing human like robots. The one of the current goals and objectives in the study of humanoid robot development is that the humanoid robot is able to walk stably both in the known environment and in an unknown environment.

A central pattern generator (CPG) is a nerve circuit in a cerebellum or a spinal cord. The CPGs generate rhythmic patterns for controlling their legs or arms. The CPGs have a feature that the generated pattern converges to the nominal pattern even if the CPG circuits are stimulated by some disturbance from the outside environment. Thus it is expected that a usage of the CPGs helps a humanoid robot to walk stably when the gait pattern is generated by the CPGs with drawing an attractor constructing by an interaction between CPGs dynamics and the environment.

From this point of view the study of the CPGs has been attracting attention in order to develop a self adaptive walking in the dynamic environment [1], [2], [3]. However an appropriate setting for walking parameters is one of the outstanding issues. To achieve a good performance in biped locomotion, the parameters of CPG model and the walking trajectory should be adjusted appropriately. This task were often done by a trial and error approach, thus it was hard to find a good setting for various environment. For this issue, Uchitane and Hatanaka have used Evolution Strategy (ES) as a tuning algorithm and showed its availability in learning a fast and stable walking pattern under the environment of RoboCup 3D Soccer Simulator [4].

Evolutionary computation is a promising approach to such parameter tuning problems, since it required only the evaluation of objective performance, for example an evaluation of the robotic locomotion. However the performance presented in [4] may not be enough. We consider the reason of this is that all parameters concerning with the CPGs, the controller and the walking trajectory are combined into one real coded vector that is an individual in the population of ES even though there are interactions between parameters, it makes difficult to search good parameters. In this paper, we consider a improvement of the speed to get a set of parameters by introducing a mask operator that is able to avoid changing partially in mutation operator in ES. This mask looks like dividing an individual into several individuals intuitively. Then, to illustrate the effect of this approach, experimental studies by using the RoboCup 3D Soccer Simulator where the actual humanoid named “Nao” made by Aldebaran Robotics is modeled are carried out.

II. WALKING STRUCTURE

A CPG is a kind of neural networks in vertebrates and it generates electric signal for controlling the movement of their body, flexing their muscle in cyclic motion for example bird’s wing stroke, bipedal walking and so on. In practice, to apply the CPG to artificial systems we use a nonlinear oscillator model as a neural oscillator. In this paper, the structure of bipedal walking of humanoid is constructed of the three components, i.e., the CPGs, foot trajectory generator and joint angle controllers. The trajectory generator gives the joint’s positions step by step based on the output of the CPG. Then, the controller has a role to let the each joint angle track its reference given by the trajectory generator. The detail of each component is described in the following.

We use the phase oscillator model proposed by Tsuchiya et.al. [5] as a CPG model. The gait and roll motions are controlled based on the output signal from phase oscillators. The dynamics of phase oscillator is described by

$$\dot{\phi}_i(t) = \omega_i + \sum_{j(j\neq i)}^{3} w_{ij} \sin\left(\phi_j(t) - \phi_i(t) + \delta\theta_{ij}\right), \qquad (1)$$

and a diagram of this oscillator network is shown in Fig.1. Where, i is a suffix to indicate the corresponding angle, i.e. $i = 1$ corresponds to rolling motion, $i = 2$ and $i = 3$ correspond to left and right legs, respectively. $\phi_i(t)$ is a phase at time instant t, ω_i and w_{ij} are frequencies

PR0001/10/0000-1458 ¥400 © 2010 SICE

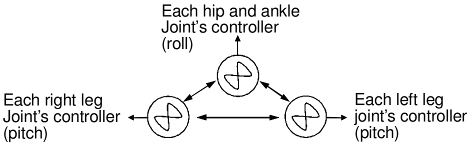

Fig. 1. Phase Oscillators.

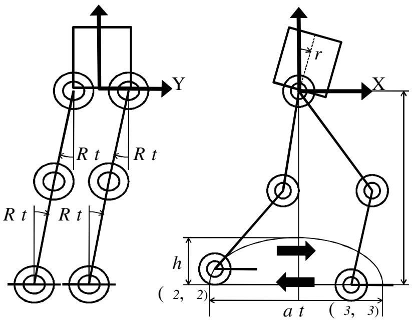

Fig. 2. Trajectory of feet and physical parameters

and connection weights between CPGs, respectively. The constraint conditions of CPG are that

$$\omega_1 = \omega_2 = \omega_3, w_{12} = w_{21},$$
$$w_{23} = w_{32}, w_{31} = w_{13}, w_{12} = w_{23}. \tag{2}$$

Now, we show a control method for each leg joint for generating the gait pattern. Let define the ankle positions $(x_i(t), z_i(t))$, where the origin is center of hip joint, for swing phase,

$$\begin{cases} x_i(t) = \alpha_t \cos(\phi_i(t)) \\ z_i(t) = -H + h\sin(\phi_i(t)) \end{cases} \tag{3}$$

and for support phase,

$$\begin{cases} x_i(t) = \alpha_t \cos(\phi_i(t)) \\ z_i(t) = -H \end{cases} \tag{4}$$

Here, $x_i(t)$ and $z_i(t)$ represent horizontal and vertical position of ankle respectively. The Fig.2 shows geometric model of humanoid legs, here each joint is connected by leg parts.

The attitude of the humanoid robot is defined by H and r at $t = 0$ and $t = 0$ means the time when the humanoid robot starts walking in this paper. The initial conditions of CPG are that

$$\phi_1(0) = 0, \phi_2(0) = \pi, \phi_3(0) = 0,$$
$$\dot{\phi}_1(0) = 0, \dot{\phi}_2(0) = 0, \dot{\phi}_3(0) = 0. \tag{5}$$

The humanoid robots bend their hip, knee and ankle joints to adjust the hight of their hip from the ground to H and to adjust the angle of forward tilt to r with satisfying condition of the following equation,

$$\theta_{hip}(t) + \theta_{knee}(t) + \theta_{ankle}(t) = 0. \tag{6}$$

By using inverse kinematics, we can derive target angles of hip, knee and ankle joints, when the hip and ankle positions are decided under the constrained condition equation (6).

TABLE I
THE LIST OF WALKING PARAMETERS

parameters name	min.	max.
$\omega_1 (= \omega_2 = \omega_3)$	2.0	10.0
$w_{12} (= w_{21} = w_{13} = w_{31})$	-1.0	1.0
$w_{23} (= w_{32})$	-1.0	1.0
α_{max}	0.0	$2\sqrt{(L1+L2)^2 - H^2}$
α_0	0.0	α_{max}
St	0.0	α_{max}
h	0.0	$\frac{L1+L2}{2}$
H	$\frac{L1+L2}{2}$	$L1 + L2$
r	0.0	30.0
$roll_{max}$	0.0	10.0
P_{gain}	0.0	5.0
D_{gain}	-5.0	5.0

Where, α_t indicates the length of the line trajectory given by the following equation,

$$\alpha_t = \min(\alpha_{t-1} + St, \alpha_{max}) \tag{7}$$

where St represents a increment of step size and α_0 means an initial step size. Then, h indicates the length of short axis of the upper half ellipse. H is the hight from grand to hip and r is the angle of forward tilt. $R(t)$ represents a roll angle at time instant t, defined by the following equation

$$R(t) = roll_{max} \sin(\phi_1(t))$$

and the $roll_{max}$ is the maximum value of roll angle given by a prior in consideration with a specification of the robot structure. Finally using PD control scheme, the angle of each leg joint (hip, knee and ankle) are controlled to track the target angle.

Table I shows the list of walking parameters and their search ranges. Hear $L1$ is the upper leg length and $L2$ is the lower leg length. In the model of "Nao," $L1$ is equal to 0.14 and $L2$ is equal to 0.11. P_{gain} and D_{gain} are the proportional gain and differential gain PD controller. As shown in Table I there are twelve parameters in order to make a good locomotion in some scene.

III. Evolution Strategy

Evolutionary Computation, e.g, Evolution Strategy (ES) and Genetic Algorithm (GA), is a powerful tool for searching the optimal solution of various optimization and/or parameter tuning problems. In this section, we introduce ES for tuning the parameters described at Section II.

Uchitane and Hatanaka applied $(\mu + \lambda)$–ES to this parameter tuning problem [7]. $(\mu + \lambda)$–ES keeps μ individuals as elite, then generates λ children from the elite individuals in each generation. ES makes children from patents by a mutation operator. In general, the mutation is performed by adding random numbers distributed with mean zero and variance σ^2 to each locus of each parents.

A fitness function should be defined in consideration of the objective of the humanoid robot walking. Uchitane and Hatanaka defined the fitness function as,

$$fitness = D - |d_l|. \tag{8}$$

Where D is equal to $\sqrt{X^2 + Y^2}$ from starting point $(0, 0)$ to ending point (X, Y) where a humanoid robot reaches without falling down. And $|d_l|$ is the Y–axis position Y.

When the robot start to walk, it directs the front of it's body to positive X–axis. This fitness value means that the humanoid robot which is able to walk fast and straight with a set of parameters in time gets a large value and the set of parameters which mark the larger fitness value is better.

Uchitane and Hatanaka showed $(\mu+\lambda)$–ES in this straightforward manner is able to give good parameter set for bipedal walking in the humanoid model [7]. However, the evolution is not enough performance, more effective approach is required. From this viewpoint, we propose a novel evolution strategy that uses a mask operator at mutation step. In this approach, all parameters are divided into several groups, then one or more groups are masked at the mutation step among predefined generations. This means a part of parameters does not change by the mutation on this generation.

In order to compare proposed ES with conventional ES, we use some common settings and a same initial population for each examination. The individuals which have a set of randomly parameters easily fall down and it is not reasonable to accept the such individual as parents population. Thus the initial population was made effectively by following way. Each individual whose parameter's values are set at random is generated then each generated individual is tested for walking. And a individual which can walk unit length or more is selected as a member of initial population. We made ten sets of forty parameters as initial populations. Then we accept the highest initial population as the initial population for the examinations in Section IV.

IV. Numerical Experiment

We use the environment of RoboCup 3D Soccer Simulation for numerical experiments. Since two libraries, "simspark" and "rcssserver3d" are published in GPL license to build a 3D soccer simulator, anyone is able to examine his/her robot motion and soccer strategies. After RoboCup 2008, Aldebaran's "Nao" is employed as the actual humanoid model in the simulator. In this study, we use simspark-0.1.2 and rcsserver3d-0.6.2. The simulator calculates movements of all humanoids and a ball in the soccer field, in consideration with their dynamics and also in consideration with the interaction between the objects including the ground. Here, it is assumed that the environment is uniform, that is, the ground is flat and the friction force is constant.

In [7], σ was randomly selected as either 0.1 or 0.33 for the mutation operator of $(\mu+\lambda)$–ES and μ and λ ware equal to 40 respectively. Here we use the same configuration.

In this study, all parameters are divided into two groups. This grouping is done for each group to have almost same number of the parameters and for parameters in the same component to be in the same group. Table II illustrates the grouping. The first group contains the parameters of CPGs and PD controller. The second group contains the parameters that decide the walking trajectory. The simulation studies are carried out to show the effect of the proposed method that uses mask to avoid mutation under the masking schedule shown in Table III. The circles mean that the parameters in the group are mutated and the crosses mean that the parameters in the group are not mutated. In the Table III, the parameters in the first group are changed at first and we call this oder proposal (n) where we call this inverse order proposal $(-n)$ in the inverse order which the parameters in the second group are changed at first.

TABLE II
Grouping of the parameters

group 1	ω_1, w_{12}, w_{23}, P_{gain}, D_{gain}
group 2	α_{max}, α_0, St, h, H, r, $roll_{max}$

TABLE III
An example for time–line of masking

generations	$1,\ldots,n$	$n+1,\ldots,2n$	$\cdots$	$\cdots$
group 1	○	×	○	×
group 2	×	○	×	○

Six examinations (proposal (1), proposal (2), proposal (10), proposal (-1), proposal (-2) and proposal (-10)) are performed. The results of proposal (1), proposal (2) and proposal (10) are shown in Fig.3, Fig.4 and Fig.5 in comparison of the conventional method. And the results of proposal (-1), proposal (-2) and proposal (-10) are shown in Fig.6, Fig.7 and Fig.8 in comparison of the conventional method. Fig.3 and Fig.6 show the best value of fitness in a single run of the simulations. Fig.4 and Fig.7 show the average value of the best individuals in 18 runs of the simulations. Fig.5 and Fig.8 show the average value of the forty elite individuals in 18 runs of the simulations.

Both the conventional method and proposed method achieved large fitness value such as over 8.0 as shown in Fig.3, Fig.4, Fig.6 and Fig.7. The proposed method gives faster evolution than the conventional method shown in Fig.3, Fig.4, Fig.6 and Fig.7. This shows that the proposed methods are effectively able to search parameters in this problem. On the contrary, masking on the second parameter group makes worse the fitness at later part of the evolution. The phenomenon gives us an insight that it is difficult to search better parameter set for the first group when the second group is fixed. Thus, we can say there is no necessity for searching the first group after such generations. Though, this phenomenon should be examined more, it is now under investigation and will be presented near future.

V. Conclusions

In this paper, we have proposed modified evolution strategy for parameter tuning in learning bipedal locomotion of the humanoid model. The numerical simulation results show the proposed method is able to achieve faster evolution than the conventional approach based on a simple evolution strategy. The results give us insight that it is not easy to tune the CPG's weights and the controller's gains for the fixed trajectory. On the other hand it seems that it is rather easy to tune the parameters of trajectory generator for the fixed parameters of CPG's weight and controller's gain. The environment of this simulation studies is uniform, an application of the proposed method to dynamic environment such as non constant friction or irregular ground is further issue.

Acknowledgment

This work was supported by, Japan Society for the Promotion of Science (JSPS) Grant-in-Aid for Scientific Research (C) 22500201.

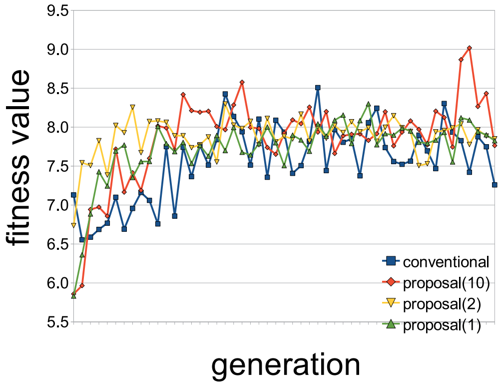

Fig. 3. This figure indicates the fitness value of the elite individual in a single run.

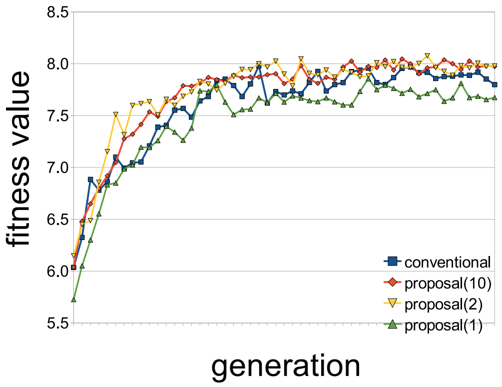

Fig. 4. This figure indicates the fitness values of the elite individual in the average over 18 run.

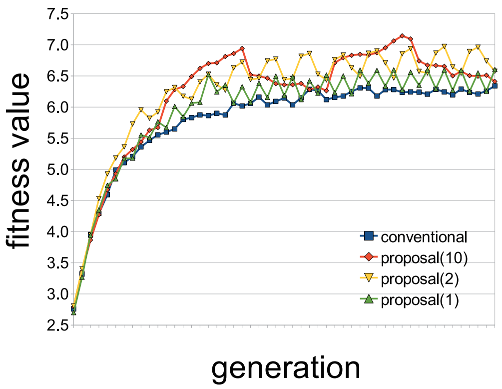

Fig. 5. This figure indicates the average fitness values of the population.

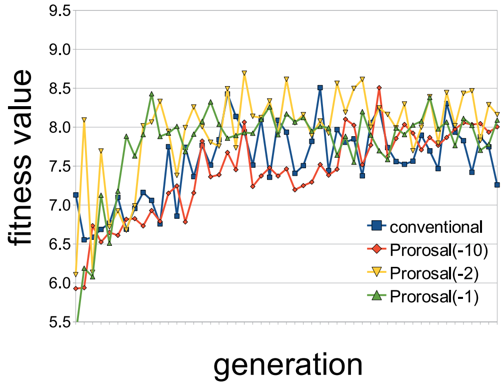

Fig. 6. This figure indicates the fitness value of the elite individual in another single run.

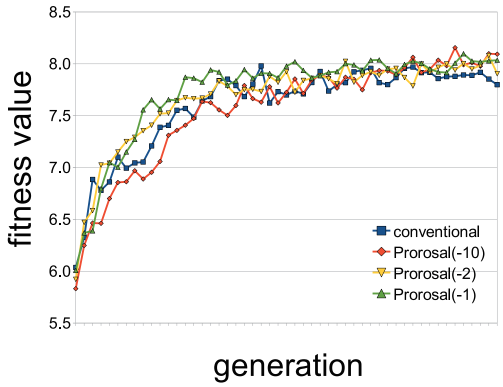

Fig. 7. This figure indicates the fitness values of the elite individual in the average over 18 run.

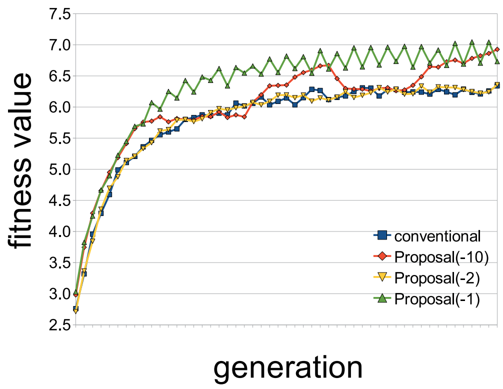

Fig. 8. This figure indicates the average fitness values of the population.

References

[1] S. Miyakoshi, G. Taga, Y. Kuniyoshi and A. Nagakubo, “Three dimensional bipedal stepping motion using neural oscillators – towards humanoid motion in the real world –,” *Proceedings of the 1998 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Vol. 1, pp.84–89, 1998.

[2] H. Kimura and Y. Fukuoka, “Biologically inspired adaptive dynamic walking in outdoor environment using a self-contained quadruped robot: ‘tekken2’,” *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 986–991, 2004.

[3] W. Yang, S. Marui, K. Murakami, W. Seo and N. Young Chong, “Adaptation in Bipedal Locomotion Using Phase Oscillator Networks,” *The 3rd International Conference on Ubiquitous Robots and Ambient Intelligence*, pp. 9–13, 2006.

[4] “rcssserver3d is a 3D soccer simulation server running on top of simspark simulation system,” http://sourceforge.net/projects/simspark/

[5] K. Tsuchiya, S. Aoi and K. Tsujita, “Locomotion control of a biped locomotion robot using nonlinear oscillators,” *Proceedings of the 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp.1745–1750, 2003.

[6] T. Bäck, U. Hammel, and H. Schwefel, “Evolutionary Computation: Comments on the History and Current State,” *IEEE Transactions on Evolutionary Computation*, Vol. 1, No. 1, pp. 3–17, 1997.

[7] T. Uchitane and T. Hatanaka, “Bipedal locomotion acquisition of humanoid agent by using evolutionary CPGs,” *Preprints of 71 th Annual Conference of Information Processing Society of Japan*, Vol.2, pp.89–90, 2009, (in Japanese).