Differentiation within Coordination in Acquisition of Skilled Throwing

Jihoon Park[1], Yuji Kawai[1], Takato Horii[1],Yuji Oshima[1],Kazuaki Tanaka[1,2], Hiroki Mori[1],Yukie Nagai[1],Takashi Takuma[3], and Minoru Asada[1]

[1] Dept. of Adaptive Machine Systems, Graduate School of Engineering,
Osaka University, Osaka, Japan

[2] CREST, Japan Science and Technology Agency

[3] Dept. of Electrical and Electronic Systems Engineering,
Osaka Institute of Technology, Osaka, Japan

robocup@er.ams.eng.osaka-u.ac.jp

Abstract

Humanoid robots have a large number of degrees of freedom (DoFs). Therefore, motor learning algorithms that explore the optimal parameters of behaviors of these robots are laborious. In contrast, it has been suggested that humans can solve this large-scale problem by coordinating most body parts strongly in the early stage of learning, and then differentiating the movements to optimize a behavior for a task. We propose that heuristic exploration through differentiation within coordination of the DoFs accelerates motor learning of humanoid robots. In this paper, we demonstrate that the process decreases iteration of trials for a throwing skill in soccer. At first, all motors related to the skill are coordinated. Thus the robot explores the optimal timing for releasing a ball in a one-dimensional search space. The DoFs are released gradually, which allows to search for the best timing of the start of each joint's movement. The experiments using a real robot showed that the exploration method was very fast and practical because the solution in a low-dimensional subspace is a good approximation of the optimum.

1 Introduction

Skilled behavior of a humanoid robot is important in the robot soccer domain. Soccer skills such as throwing, kicking, and biped locomotion require coordination of the whole body with a large number of degrees of freedom (DoFs). Thus, we must solve a large-scale problem for designing a skilled behavior of a humanoid robot.

There exists much literature on the heuristic exploration approach to solve the problem. In this approach, a robot autonomously optimizes parameters of skills through practice. It has been reported that evolutionary computation (e.g., [1, 2]) and particle swarm optimization (e.g., [3, 4]) enable the robot to acquire faster gait.

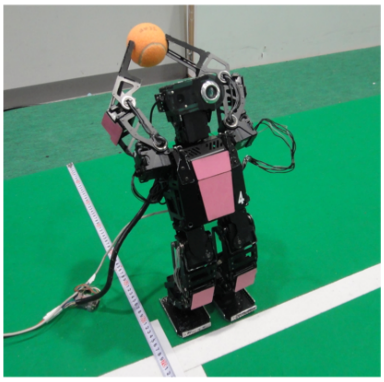


Figure 1: Throwing for exploration of optimal parameters.

Main optimization parameters in most of these studies have been trajectories of limbs or Central Pattern Generators. However, the number of iterations including the evaluation of performance was very large because the search space was vast. Also, degradation and malfunction are problems with using a real robot for a long time. Therefore, optimization methods that require many trials are not useful.

In contrast, Kohl and his colleagues [5–7] have demonstrated that Hill Climbing and Policy Gradient algorithms allow to optimize the parameters for quadruped locomotion and kicking a ball. These algorithms generally require less iterations than evolutionary computation and particle swarm optimization. However, the intrinsic complexity of a humanoid robot's body still causes a large number of iterations.

We take a hint from the progression of skills in humans during behavior acquisition in high-dimensional motor space. Bernstein [8] (see also [9, 10]) suggested freezing and freeing of DoFs in skill acquisition. In the early stage of learning of a motor skill, some DoFs are reduced (frozen). These DoFs are then released (freed) gradually as the learning progresses. These stages of motor learning allow to reduce the search space dimensionality. Yamamoto and Fujinami [11] also found a common organization of acquisition of a periodic skill: differentiation within coordination. They compared clay kneading movements for pottery of experienced subjects and experts. While the experienced subjects tend to coordinate their body parts, slight phase differences between body parts are observed in experts' movements. Their group [12] found similar results for the proficiency of samba dance. A possible interpretation of the coordination of movement in the early learning is that less movement parameters simplify the optimization for the skill.

We introduce this idea of differentiation within coordination to optimization methods, and then aim to apply this to a soccer throwing skill. Most important in the acquisition of skilled throwing is the timing of releasing the ball. A humanoid robot searches for the best timing of the start of each joint based on timing of releasing a ball through practice as shown in Fig. 1. All joints related to the throwing skill are initially coordinated. That is, the robot roughly optimizes the timing of releasing a ball in a one-dimensional space. The joints are then gradually released, which allows the robot to search more optimal parameters. As a result, the robot acquires skilled throwing with only a small number of evaluations.

This paper is organized as follows: In section 2, we explain heuristic exploration using differentiation within coordination of DoFs. Throwing parametrization of a robot and the experimental setting are described in section 3 and 4, respectively. Section 5 then demonstrates that the proposed optimization method results in quicker exploration of optimal parameters. In section 6, the results are discussed, and in section 7 we conclude our research.

2 Heuristic exploration through differentiation within coordination

2.1 Differentiation within coordination

An example of exploration through differentiation within coordination of parameters is shown in Fig. 2. Here, we assume a two-dimensional search space, that is, the only parameters are S_1 and S_2. There are two stages of optimization: coordination and differentiation.

2.1.1 Coordination

The search space is restricted to coordination of all parameters, namely, $S_1 = S_2$. An initial value is selected in this one-dimensional space (on the dashed line in Fig. 2). The optimal parameter is then explored on the line.

2.1.2 Differentiation

The restriction is gradually lifted after finishing the optimization in the previous search space. The search space is hence extended to multiple dimensions. Initial values are the best parameters in the previous stage. A solution of this algorithm is the optimal set of parameters after all parameters have been freed.

It is expected that this constraint on dimensionality reduces the search space, which accelerates the optimization.

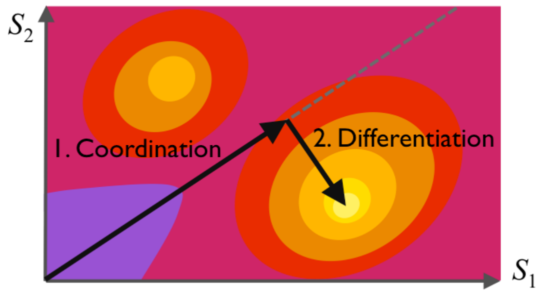


Figure 2: A conceptual example of the proposed exploration. In case of a two-dimensional objective function, the two parameters, S_1 and S_2, are coordinated at first. These parameters are differentiated after optimization in the one-dimensional space. The constraint on the search space enables faster exploration.

2.2 Optimization method

Every time the dimension increases, an optimization method must be applied to the search space. This paper uses a hill climbing algorithm and a modified particle swarm optimization (PSO). These algorithms are widely applied to parameter optimization problems (see [3–5, 7]).

2.2.1 Hill climbing

The Hill Climbing algorithm is one of the simplest optimization methods. It is known that this algorithm explores a solution quickly. An initial value is selected and evaluated in the search space. All neighbors of the initial parameter are evaluated, and the highest-scoring parameter among the neighbors is selected. The selected value is the next center, and then its neighbors' values are evaluated. Repeat the evaluation and the selection until no higher score can be found.

2.2.2 Modified particle swarm optimization

PSO [13] is a probabilistic optimization method similar to genetic algorithms. Initially, a swarm of N particles is generated in the D-dimensional search space. Here, we introduce an initial value to this algorithm so that the optimization can inherit the best parameter of the previous search space. Although the existing PSO initializes the particles with random values, this paper sets the initial positions according to a normal distribution, where its mean and variance are the initial value and $\mathbf{v}$, respectively. These particles are assigned a position $\mathbf{x_i}$ $(1 \le i \le N)$ and a velocity $\mathbf{v_i}$ $(1 \le i \le N)$, which are D-dimensional vectors. Each particle is evaluated by the performance of its parameters. At each iteration, the velocity of each particle is updated depending on two values: the personal best position $\mathbf{p}best_i$ $(1 \le i \le N)$ and the global best position $\mathbf{g}best$. $\mathbf{p}best_i$ is the best position that each particle has ever evaluated. $\mathbf{g}best$ is the best position that all particles have evaluated. Each velocity $\mathbf{v}_i^t$ at iteration t is updated by:

$$\mathbf{v}_i^{t+1} = w\mathbf{v}_i^t + c_p r_p^t \times (\mathbf{p}best_i^t - \mathbf{x}_i^t) + c_g r_g^t \times (\mathbf{g}best^t - \mathbf{x}_i^t), \quad (1)$$

where, w, c_p and c_g are weights. r_p and r_g are normal random numbers between 0 and 1. We restrict the range of velocity between $-\mathbf{v}^{max}$ and $\mathbf{v}^{max}$, which is determined by:

$$\mathbf{v}^{max} = k \times \mathbf{x}^{max}, \tag{2}$$

where, x^{max} is the range of exploration in each dimension, and $0.1 \leq k \leq 1$. The next positions of particles $\mathbf{x}_i^{t+1}$ are calculated by:

$$\mathbf{x}_i^{t+1} = \mathbf{x}_i^t + \mathbf{v}_i^{t+1}. \tag{3}$$

We stop exploring when $\mathbf{g}best$ does not change during n iterations.

3 Throwing parametrization

For our task, a robot searches the optimized combination of the start timing of each joint to throw the ball as far as possible. The VisiON 4G robot a commercial humanoid robot manufactured by Vstone Co.,Ltd. (see Fig. 1) was used for this experiment. Fig. 3 illustrates the robot's DoFs configuration. The robot has 22 DoFs and each joint is actuated by a VS-SV410 servomotor.

From these 22 DoFs, we selected 4 DoFs for throwing:

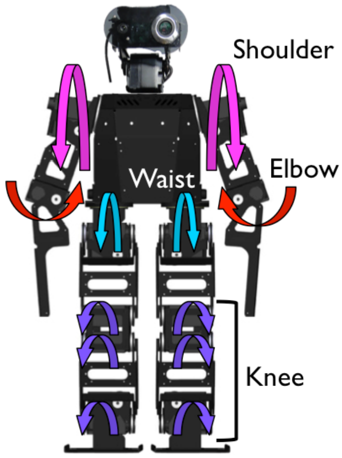
- Pitch shoulder: throwing the ball overhead.
- Roll elbow: holding and releasing the ball.
- Pitch waist: achieving more force by the reaction.
- Knee: stretching both knees, which consist of 3 motors each.

Fig. 4 shows the definition of the parameters. We did not use velocities or positions of each Degree of Freedom(DoF) but the timing of movements of each DoF as parameters. In this case, the DoF of the elbow serves as the base of timing because it directly determines the release of the ball and is therefore most important for optimizing skilled throwing. We defined the timing of the start of movement of the shoulder, knee and waist based on the elbow's timing as t_s, t_w and t_k, respectively. The robot learns the optimal $\mathbf{t} = (t_s, t_w, t_k)$ through practice. Initially, the 3 DoFs are coordinated, i.e., $t_s = t_w = t_k = t_{init}$, and the robot optimizes t_{init} (see Fig. 4(a)). In the last stage of learning, all DoFs are differentiated. Thus the robot searches the optimal $\mathbf{t}$ in the three-dimensional space (see Fig. 4(b)).

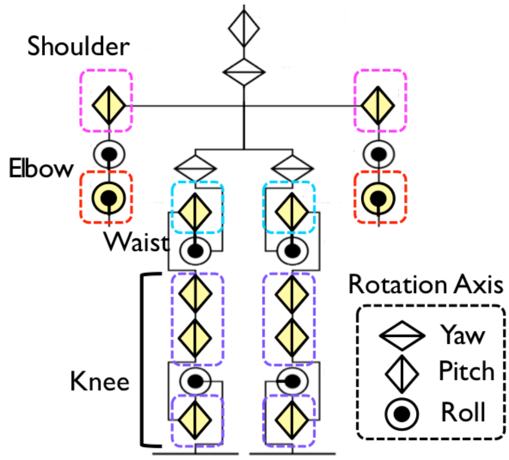
4 Experimental setting

In order to validate the proposed optimization method, we conducted experiments using a real robot. The robot explores optimal combinations of t_s, t_w, and t_k.

The evaluation was determined by the distance between a robot's toe and the ball point. The throwing distance was measured by visual inspection through video recording with a measuring tape as shown in Fig. 5. We evaluated the distance of throwing regardless of whether the robot fell down during the trials.



(a) Front view of VisiON 4G and its essential DoFs for throwing.

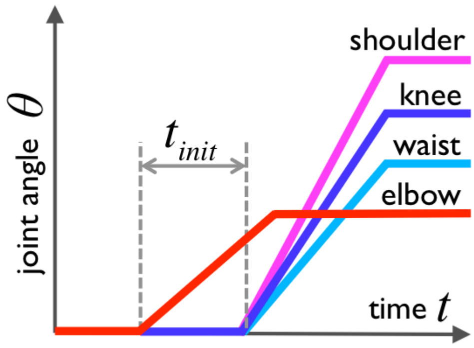


(b) DoFs configuration of VisiON 4G. The yellow joints are used in this experiment.

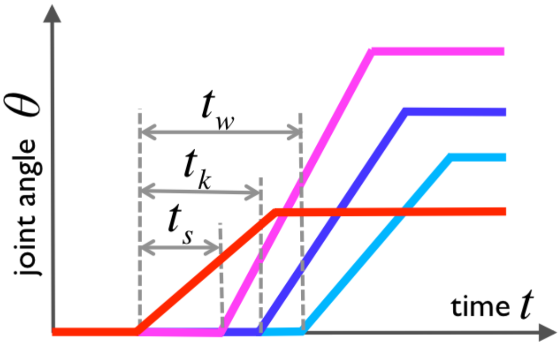
Figure 3: The number of the substantial DoFs used in the current experiment is 4: the pitch shoulder, the roll elbow, the pitch waist and the pitch knee. We assume symmetry of the motors. The DoF of the knee consists of 6 motors. The elbow affects the holding and releasing the ball.

The robot's motion started from the same initial pose as shown in Fig. 1 in each trial. We gave the ball to the robot so that the robot could hold the ball with both hands. It took 10 steps to execute a throwing motion, where 1 step was 1/30 sec. The range of exploration was set to [-5, +2] based on the start timing of the elbow, and the range segmentation was 1 step. The robot rested for 5 minutes after every 10 trials to prevent overheating of the motors.

Optimization experiments were conducted off-line, and were applied to the dataset obtained by exhaustive search in advance. Two trials of the experiment were performed, each of which consisted of 512 different timings tested. The objective function was given the mean of two trials. We tested four optimization algorithms: Hill Climbing and PSO through differentiation within coordination, and existing Hill Climbing and PSO. We then compared the number of evaluations and achieved



(a) Coordination of DoFs in the first stage



(b) Differentiation of DoFs in the latter stage

Figure 4: The number of parameters increases gradually during learning. In the early stage of learning (a), the DoF of the shoulder, the knee and the waist are coordinated. The robot explores the optimal t_{init}, namely timing of releasing of the ball in the one-dimensional space. In the last stage of learning (b), the timing of the start of the each DoF, t_s, t_w, t_k, is optimized.

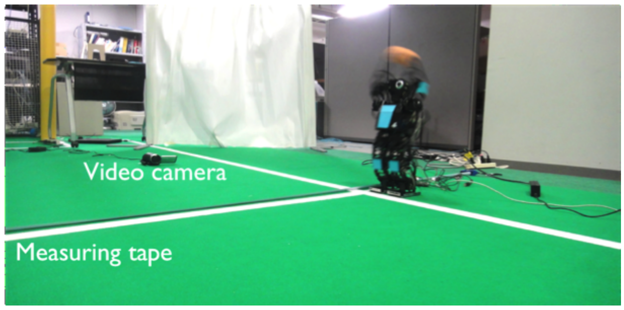
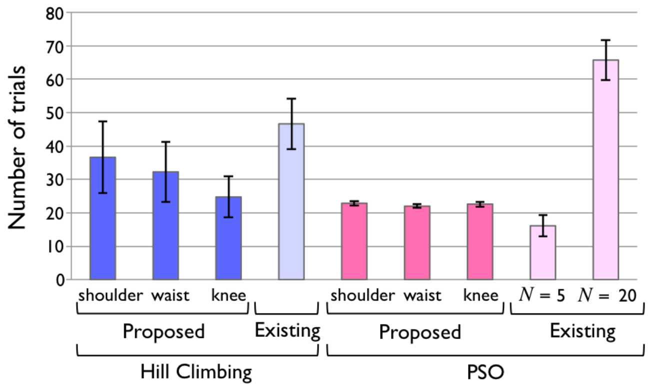
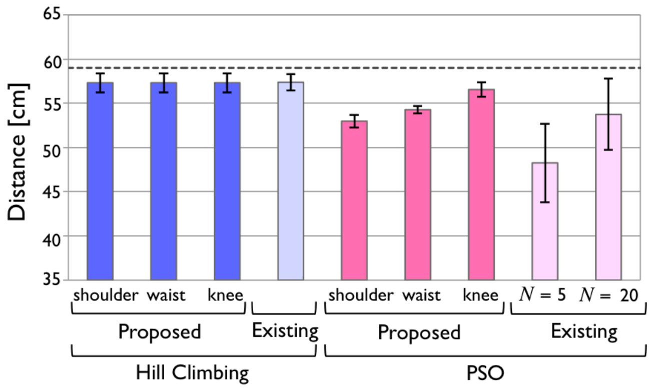


Figure 5: The experimental environment to optimize the parameters for throwing. We record the distance between a robot's toe and a ball fall point.



(a) The number of trials



(b) Throwing performance. The dashed line denotes the global optimum (59 cm).

Figure 6: The results of each optimization method. The blue and the red bars indicate the results of the Hill Climbing algorithm and PSO, respectively. The proposed methods differentiation of DoF of shoulder (left), waist (middle), knee (right) from other DoFs in the second stage. N is the number of particles.

optimal performance.

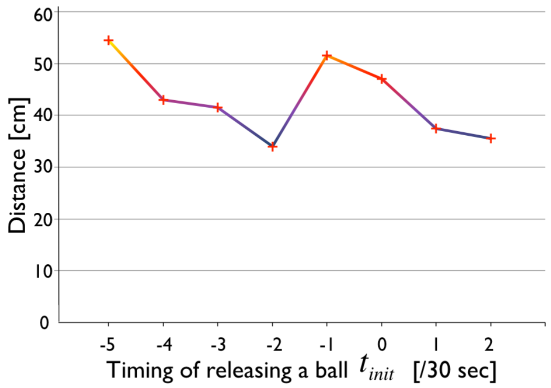
In the Hill Climbing algorithm, one iteration needs 26 evaluations in the three-dimensional search space. However, we did not count the evaluations of the parameters where the robot once searched. The variables in the PSO were empirically determined: 5 particles were initially positioned according to a normal distribution, whose variance was set to 3. We set $w = c_p = c_g = 0.5$ in Eq. (1) and $k = 0.25$ in Eq. (2). The optimization was finished when $\mathbf{g}_{best}$ did not change for 3 iterations.

The initial parameter was given as an integer between -5 and 2 (i.e., 8 patterns). Each optimization method was conducted 8 times for all initial parameters. The proposed PSO was ran 10 times with each initial parameter setting because PSO includes randomness. In the existing PSO, randomly-selected initial parameters are given, and then we tested it 80 times.

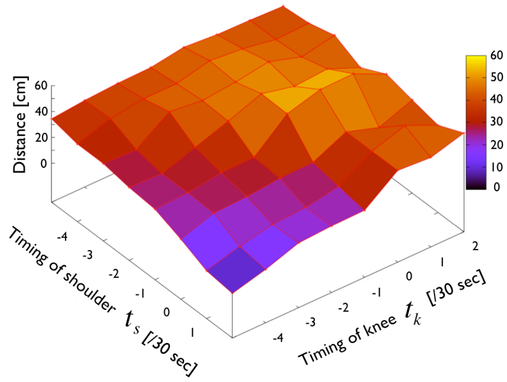
5 Result

5.1 Number of trials and throwing performance

Fig.6(a) shows the average number of trials for each optimization method. Less trials mean faster exploration, which relieves robot of load. The blue bars at the left side and red bars at the right side denote the results of the Hill Climbing algorithm and PSO, respectively. There are three results in the proposed optimization through differentiation within coordination: shoulder (left), waist (middle) or knee (right) were differentiated from other DoFs in the second stage. We can see that both proposed methods result in less trials than the existing methods. However, the result of PSO using 5 particles ($N = 5$) shows the least number of trials.



(a) one-dimensional search space



(b) two-dimensional search space

Figure 7: Objective function. Proposed algorithm optimizes t_{init} in the one-dimensional space (a) in the first stage of exploration. In two-dimensional search space with $t_w = -1$ (b), the global optimal parameter $\mathbf{t}^{opt}$ is (-1, -1, 1) which results in a distance travelled of 59cm.

The optimized performance of each method is shown in Fig.6(b). The dashed line indicates the global optimum. We find that all results of Hill Climbing show high performance. The results of the proposed PSO through coordination show less variance and have nearly the same performance as the existing PSO with $N = 20$. The existing PSO with $N = 5$ needs the least number of trials but also performs worst. Therefore, the proposed method can reduce the number of trials while maintaining the high performance of the existing PSO.

5.2 Objective function

As shown in Fig. 7, the objective function is obtained by exhaustive search in order to reveal the cause of the above result. Fig. 7(a) illustrates a one-dimensional objective function, where the optimal t_{init} values are explored in the first stage. There are two local maxima: $t_{init} = -1, -5$. The global optimum $\mathbf{t}^{opt}$ is (t_s, t_w, t_k) = (-1, -1, 1) as shown in Fig. 7(b). Thus, the result of optimization in the one-dimensional space should be -1 so that the robot can finally find $\mathbf{t}^{opt}$. In Hill Climbing, the optimal t_{init} is -1 if the initial value is more than -2. All particles, however, move toward -5 in PSO. This is why the optimization by PSO with coordination of DoFs was worse than by Hill Climbing (see Fig. 6(b)).

The coordination of the DoF of the shoulder and the waist makes it easier to reach $\mathbf{t}^{opt}$ because $\mathbf{t}^{opt}$ is $t_s = t_w = -1$. Thus, the optimization through coordination of the DoF of t_s and t_w in the second stage results in the best performance. The proper order of releasing the DoF may be task-dependent.

6 Discussion

It is hard for a robot to acquire skilled throwing. Even if we use a faithful simulator or make a dynamical mathematical model of a robot, there exists a gap between the real and the virtual world. One of the differences originates from the in environmental complexity. The robot's body interacts with the ball throwing. The ball deforms slightly and the robot undergoes reaction forces. This interaction seems to influence the performance. Most simulators, however, cannot address detailed touch calculations. The inherent delay of motors from motor commands is also a crucial problem. Many athletic behaviors such as throwing are instantaneous movements. The throwing took only 1/3 sec in this experiment. Thus the motor's tiny delay makes a difference of performance. After all, it is necessary for acquisition of skilled behavior to optimize in high-dimensional space using a real robot.

We demonstrated that the high dimensionality may be reduced by differentiation within coordination of the DoFs. In particular, this optimization method might be applicable for athletic behaviors. Many body parts move instantaneously in parallel moment, which can be regard as coordination of body parts. However, a little differentiation of movements of each body part is observed in human's skilled motion. In the throwing by an expert the timing to maximum velocities of body parts does not always correspond to the timing of releasing of a ball [14]. It means that there are tiny phase displacements between DoFs for effective movements. The timing optimized in coordination of the DoFs is close to the global optimum. Thus, the local maximum reached in the first stage of exploration is useful even if the space dimensionality increases.

Our proposed optimization method can be applied to other skills. A slight differentiation of the timing of a leg's DoF may be important in high-kicking (kicking the ball as high as possible), which has been an official technical challenge in the RoboCup soccer humanoid league since 2012. We will attempt to optimize these soccer skills by the proposed method. In addition, velocity of body parts is also important for skilled behavior. We will address the acquisition of skilled behavior with more parameters.

7 Conclusion

In this paper we presented a practical optimization method through differentiation within coordination of a robot's body parts. All the DoF related to the skill were coordinated in the first stage of learning. Thus, the robot optimized the timing of the start of releasing the ball in one-dimensional space. The DoF were then differentiated one by one, which enabled the robot to explore the optimal timing of the start of each joint's movement. The reduction of the search space dimensionality, consequently, could decrease the number of trials. Our experiments showed that the optimization through coordination of the DoF resulted in as high performance as the result of optimizing without coordination even if less trials were used.

This optimization method may be leveraged when acquiring quick movements such as throwing, kicking and so on. Instantaneous athletic skills can be coordinated behaviors. Thus the optimization of coordination of DoF might be more plausible, i.e., not just a local solution. The robot can reach quickly a valid solution because of usage of the best solution in the previous stage.

Acknowledgment

The authors gratefully acknowledge the contribution of the team members of the JST ERATO Asada Project(JEAP) for all materials, and also Associate Prof. Tomomichi Sugihara and a former student, Hirotaka Kimura (Currently, the City of Nagoya, Japan) at Osaka University for their suggestion about heuristic exploration.

References

[1] GS Hornby, M. Fujita, S. Takamura, T. Yamamoto, and O. Hanagata. Autonomous evolution of gaits with the sony quadruped robot. In *Proc. of the Genetic and Evolutionary Computation Conference*, volume 2, pages 1297–1304, 1999.

[2] G. Daoxiong, Y. Jie, and Z. Guoyu. A review of gait optimization based on evolutionary computation. *Applied Computational Intelligence and Soft Computing*, 2010, 2010.

[3] C. Rong, Q. Wang, Y. Huang, G. Xie, and L. Wang. Autonomous evolution of high-speed quadruped gaits using particle swarm optimization. *RoboCup 2008: Robot Soccer World Cup XII*, pages 259–270, 2009.

[4] N. Shafii, S. Aslani, O. Nezami, and S. Shiry. Evolution of biped walking using truncated fourier series and particle swarm optimization. *RoboCup 2009: Robot Soccer World Cup XIII*, pages 344–354, 2010.

[5] N. Kohl and P. Stone. Machine learning for fast quadrupedal locomotion. In *Proc. of the 19th National Conf. on Artificial Intelligence*, pages 611–616, 2004.

[6] M. Saggar, T. D'Silva, N. Kohl, and P. Stone. Autonomous learning of stable quadruped locomotion. *RoboCup 2006: Robot Soccer World Cup X*, pages 98–109, 2007.

[7] M. Hausknecht and P. Stone. Learning powerful kicks on the aibo ers-7: The quest for a striker. *RoboCup 2010: Robot Soccer World Cup XIV*, pages 254–265, 2011.

[8] N.A. Bernstein. The co-ordination and regulation of movements. 1967.

[9] K.M. Newell and D.E. Vaillancourt. Dimensional change in motor learning. *Human Movement Science*, 20(4-5):695–715, 2001.

[10] B. Vereijken, R.E.A. van Emmerik, HTA Whiting, and K.M. Newell. Free(z)ing degrees of freedom in skill acquisition. *Journal of Motor Behavior*, 24(1):133–142, 1992.

[11] T. Yamamoto and T. Fujinami. Hierarchical organization of the coordinative structure of the skill of clay kneading. *Human movement science*, 27(5):812–822, 2008.

[12] K. Matsumura, T. Yamamoto, and T. Fujinami. A study of samba dance using acceleration sensors. In *Proc. of the 8th Motor Control and Human Skill Conference*, pages 5–4, 2007.

[13] J. Kennedy and R. Eberhart. Particle swarm optimization. In *Proc. of the IEEE International Conference on Neural Networks*, volume 4, pages 1942–1948, 1995.

[14] T. Reilly. *Science and Soccer*. Routledge, 1995.