Throwing Skill Optimization through Synchronization and Desynchronization of Degree of Freedom

Yuji Kawai[*1], Jihoon Park[*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 by such robots which explore the optimal parameters of behaviors is one of the most serious issues in humanoid robotics. In contrast, it has been suggested that humans can solve such a problem by synchronizing many body parts in the early stage of learning, and then desynchronizing their movements to optimize a behavior for a task. This is called as "Freeze and Release." We hypothesize that heuristic exploration through synchronization and desynchronization of DoFs accelerates motor learning of humanoid robots. In this paper, we applied this heuristic to a throwing skill learning in soccer. First, all motors related to the skill are actuated in a synchronized manner, thus the robot explores optimal timing of releasing a ball in one-dimensional search space. The DoFs are released gradually, which allows to search for the best timing to actuate the motors of all joints. The real robot experiments showed that the exploration method was fast and practical because the solution in low-dimensional subspace was approximately optimum.

1 Introduction

Skilled behaviors of a humanoid robot are important in the robot soccer domain. Soccer skills such as throwing, kicking, and biped locomotion require coordination of the whole body movements with a large number of degrees of freedom (DoFs). Designing a skilled behavior of a humanoid robot with high DoFs is one of the most serious issues.

There exist many studies on the heuristic exploration approach to solve such problems. Among them, evolutionary computation (e.g., [1,2]) and particle swarm optimization (e.g., [3,4]) enabled the robot to acquire faster gait. Main optimization parameters in these studies have been trajectories of limbs or parameters of Central Pattern Generators. However, the number of iterations including evaluation of performance was very large because of a vast. Generally, real robots are prone to be easily broken, therefore optimization methods with much less trials are desired.

Peter and his colleagues [5–7] have demonstrated that Hill Climbing and Policy Gradient algorithms successfully optimized the parameters for quadruped locomotion

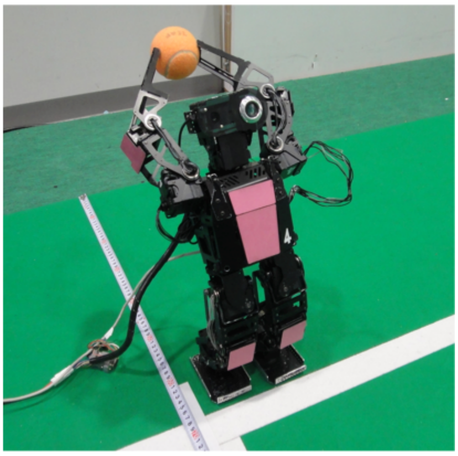


Fig. 1. Throwing for exploration of optimal parameters.

and kicking a ball. These algorithms converge to solution more rapidly than evolutionary computation and particle swarm optimization. However, the complexity of a robot's body still intrinsically causes a large number of iterations.

We take an idea from the progression of skills in humans who have their high-dimensional motor space. Bernstein [8] (see also [9, 10]) suggested freezing and releasing of DoFs in skill acquisition. In the early stage of learning of a motor skill, some DoFs are reduced (frozen). The DoFs are then released (freed) gradually as the skill progresses. These stages of motor learning enable to reduce the search space dimensionality. Yamamoto and Fujinami [11] also found a common organization of acquisition of a periodic skill: synchronization and desynchronization. They compared clay kneading movements for pottery of experienced subjects with those of the experts. While the experienced subjects tend to synchronize their body parts, slight phase differences between body parts are observed in the experts' movements. Their group [12] found the similar process for proficiency of samba dance. A possible interpretation of the synchronization of movements in the early learning for the skill is that smaller number of parameters of the movements simplify the optimization by reducing the dimensionality.

We introduce this idea of synchronization and desynchronization to optimization methods, and then apply the method to progress of soccer throwing skill. Of importance in the acquisition of skilled throwing is the timing when to release the ball. A robot searches for the best timing to actuate each joint based on timing of releasing a ball through practice as shown in Fig. 1. All joints related to the throwing skill are synchronized initially. That is, the robot optimizes roughly the timing of releasing a ball in 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 even with a small number of trials.

This paper is organized as follows: In section 2, we explain heuristic exploration using synchronization and desynchronization of DoFs. Throwing parametrization of a

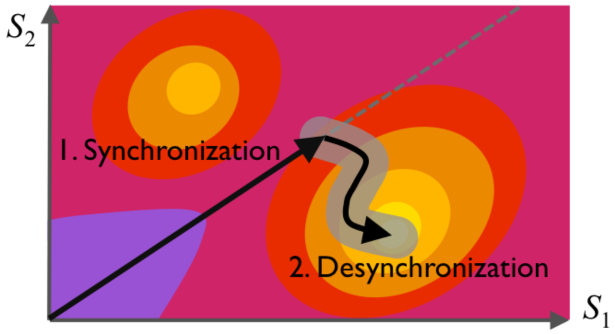


Fig. 2. An example of exploration by Hill Climbing algorithm through synchronization and desynchronization. In case of a two-dimensional objective function, the two parameters, S_1 and S_2, are synchronized at first. These parameters are desynchronized after optimization in one-dimensional space. The glay area means the neighborhood of a local solution in two-dimensional space. The constraint on the search space enables faster exploration.

humanoid robot and the experimental setting are described in sections 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 given and future issues are discussed, and in section 7, we conclude our research.

2 Heuristic exploration through synchronization and desynchronization

2.1 Synchronization and desynchronization

Fig. 2 illustrates an example of exploration through synchronization and desynchronization of parameters. Here, we assume a two-dimensional search space, that is, the only two parameters are S_1 and S_2. There are two stages of optimization: synchronization and desynchronization.

Synchronization The search space is restricted to synchronization of all parameters, namely, $S_1 = S_2$ $(= S_3 = \cdots = S_n)$. An initial value is selected in one-dimensional search space (on the dashed line in Fig. 2). An optimal parameter is then explored on this line.

Desynchronization The restriction is gradually lifted after finishing the optimization in the previous search space. The search space is hence extended to one of other dimensions. Initial values are the best one in the previous stage. A solution of this algorithm is an optimal set of parameters when all parameters are explored.

The following is a procedure:

1. Synchronization process (above): one-dimensional search by freezing.

2. $i = 1$, and repeat the following until all dimensions are explored.
   2-1 Release one dimension ($i = i + 1$) and apply an optimization method starting from the optimal solution in the previous stage (before releasing) as the initial value in i-dimensional search space.
   2-2 Find the optimal one in this space. If $i = n$ (the full dimension), then the optimal one is globally optimal. Else, go to 2-1

Therefore, the dimension of the search space is gradually increasing while the search area is expected to decreasing owing to starting the optimization from reasonable initial value.

2.2 Optimization method

Every time one dimension is released, an optimization method is applied in the search space. We use a Hill Climbing algorithm and a modified particle swarm optimization (PSO). These algorithms are widely applied to parameter optimization problems (see [3–5, 7]). In the both algorithms, the initial value in next search space is solution in previous one.

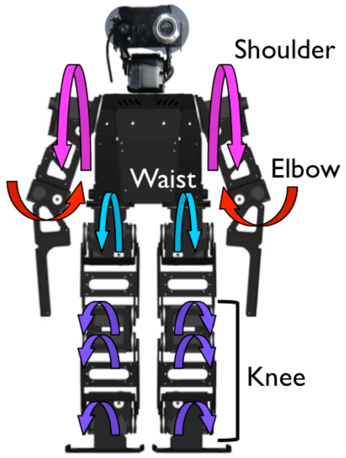
Hill Climbing A Hill Climbing algorithm is one of the simplest optimization methods. It is well-known that this algorithm explores a solution quickly. An initial value is selected and evaluated in the search space. All neighbors of the initial one are evaluated, and the highest-scoring parameter among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

Modified particle swarm optimization (PSO) PSO [13] is a probabilistic optimization method just as 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 parameters in the previous search space. Although the existing PSO gives randomly the positions of initial particles, we give the initial positions according to normal distribution, where its mean and variance are the initial value and v, respectively. These particles are assigned a position $\mathbf{x_i}$ and a velocity $\mathbf{v_i}$ ($1 \leq i \leq N$), which are both D-dimensional vectors. Each particle is evaluated by the performance. At each iteration, the velocity of each particle is updated depending on two values: the personal best position $\mathbf{p}best_i$ ($1 \leq i \leq N$) 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:

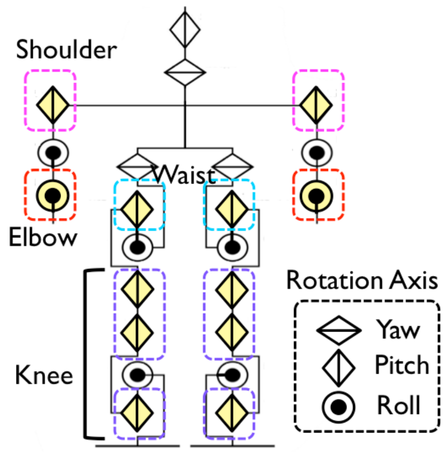
$$\mathbf{v}_i^{t+1} = w\mathbf{v}_i^t + c_p r_p^t \times (\mathbf{p}best_i^t - \mathbf{x}_i^t) + c_g r_g^t \times (\mathbf{g}best^t - \mathbf{x}_i^t), \tag{1}$$

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:

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



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



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

Fig. 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.

where, $\mathbf{x}^{max}$ 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 judge the end of the exploration when $\mathbf{g}best$ does not change during n iterations in the current experiment.

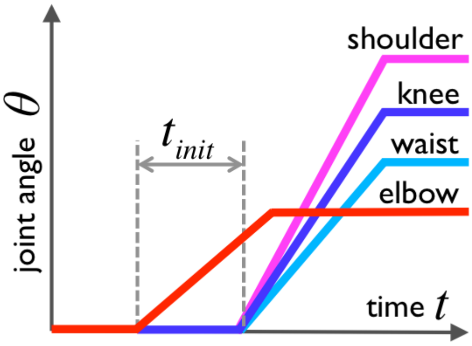
3 Throwing parametrization

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 was used for this experiment (see Fig. 1), a commercial humanoid robot manufactured by Vstone Co.,Ltd. Fig. 3 depicts the robot's DoF configuration. The robot has 22 DoFs and each joint is actuated by a VS-SV410 servomotor.

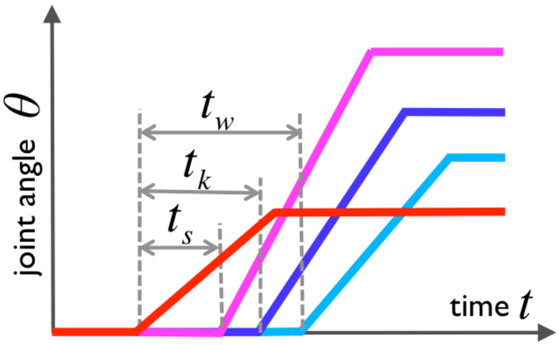
We, however, selected essential 4 DoFs for throwing:

- Pitch shoulder: throwing the ball overhead.
- Roll elbow: holding and releasing the ball.
- Pitch waist: achieving more force by the reaction.
- Knee: stretching the both knees, which consisting of 6 motors.

Unfortunately, VisiON 4G does not have the DoF of pitch elbow, which is required for human throwing.



(a) Synchronization of DoFs in the first stage



(b) Desynchronization of DoFs in the latter stage

Fig. 4. The number of parameters increases gradually in the learning. In the early stage of learning (a), the DoFs of the shoulder, the knee and the waist are synchronized. 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 and t_k, is optimized.

Fig. 4 shows the definition of the parameters. We did not use velocities or positions of individual DoFs but the timing of movements of three DoFs (shoulder, knee, and waist) as the parameters. Here, DoF of the elbow is a base of the timing because it is important for a skilled throwing to optimize the timing of releasing the ball. We defined the timing of the start of movements 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 synchronized, i.e., $t_s = t_w = t_k = t_{init}$, and then the robot optimizes t_{init} (see Fig. 4(a)). Secondly, one DoF, the shoulder, the waist or the knee, is differentiated from other DoFs. If the DoF of the shoulder is selected here, the parameters are t_s and $t_w = t_k$. In the last stage of learning, all of DoFs are desynchronized. Thus the robot searches optimal $\mathbf{t}$ in the three-dimensional space (see Fig. 4(b)).

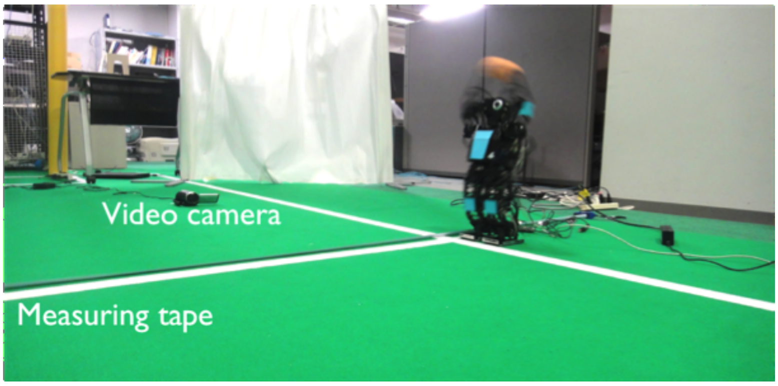


Fig. 5. The experimental environment to optimize the parameters for throwing. We recorded the distance between a robot's toe and a ball fall point.

4 Experimental setting

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

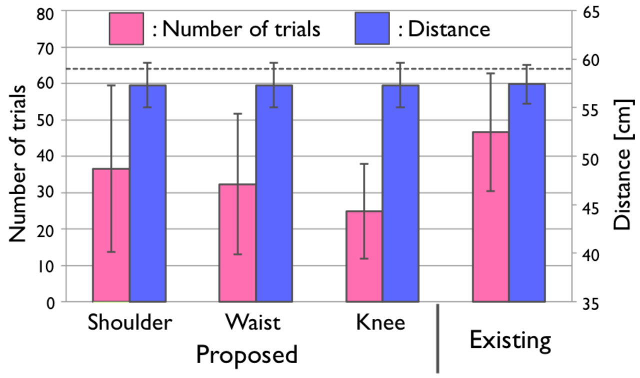
The performance is evaluated by the distance between a robot's toe and a ball fall point. The throwing distance is measured by a visual inspection through video recording with a measuring tape as shown in Fig. 5. We evaluate the distance of throwing regardless of the posture after throwing (keep standing or not).

The robot starts its motion from the same initial pose as shown in Fig. 1 in each trial. We give the ball to the robot so that the robot can hold the ball with both hands. It takes 10 steps to execute throwing motion, where 1 step is 1/30 sec. The range of exploration is set to [-5, +2] based on start timing of the elbow. The robot rests for 5 minutes every time after 10 trials to prevent overheat of the motors.

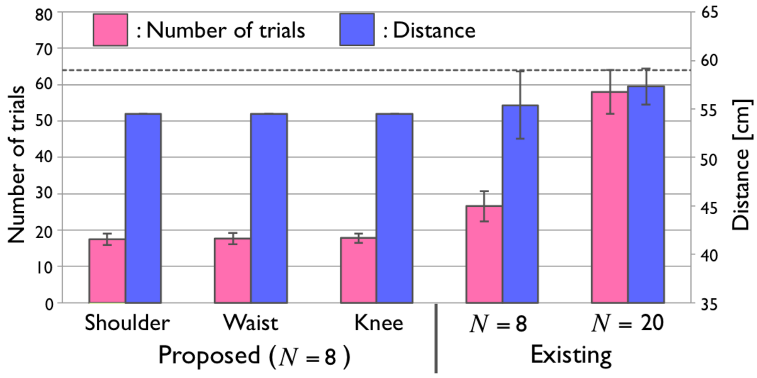
Optimization experiments are conducted off-line. We evaluate optimization methods using dataset obtained by exhaustive search in advance. Two trials of the experiment are performed, each of which consists of 512 different timings. The objective function is given the mean of two trials. We tested four optimization algorithms: Hill Climbing and PSO through synchronization and desynchronization, and existing Hill Climbing and PSO. We then compare the number of the evaluations and achieved optimal performance.

In the Hill Climbing algorithm, one iteration needs 26 evaluations in three-dimensional search space. We, however, does not count the evaluations of the parameters where the robot once searched. The variables in the PSO are empirically determined: 8 particles are initially positioned according to a normal distribution, whose variance is set to 3. We set $w = c_p = c_g = 0.5$ in Eq. (1) and $k = 0.25$ in Eq. (2). The optimization is finished when the $\mathbf{g}_{best}$ does not change for 3 iterations.

An initial parameter is given as an integer between -5 and 2 (i.e., 8 patterns). Each optimization method is conducted 8 times for all initial parameters. The proposed PSO



(a) Hill Climbing algorithm



(b) PSO

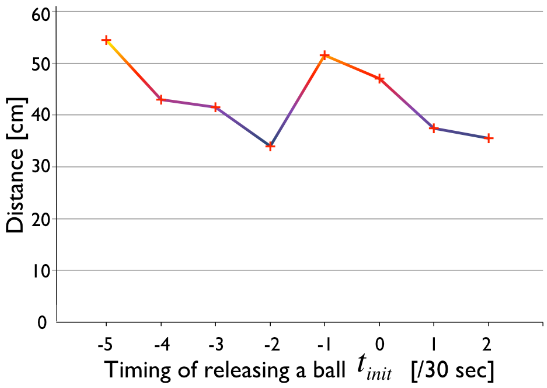
Fig. 6. The results of (a) Hill Climbing algorithm and (b) PSO. The blue and the pink bars indicate the number of trials and throwing performance, respectively. The proposed methods desynchronized DoF of shoulder (left), waist (middle), or knee (right) from other DoFs in the second stage. N is the number of particles. The dashed line denotes the global optimum (59cm). Each errer bar indicates the standard deviation.

is 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 test it 80 times.

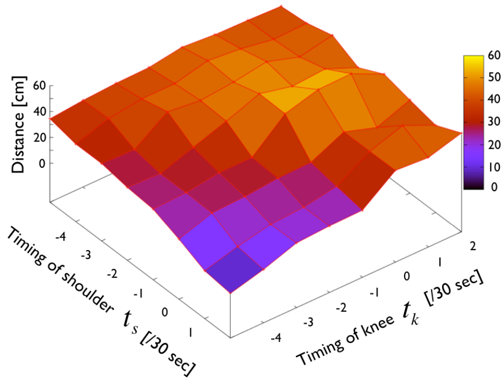
5 Result

5.1 Number of trials and throwing performance

Fig. 6 shows the results of optimization methods: (a) Hill Climbing algorithm and (b) PSO. The pink bars denote average of the number of trials in each optimization method. Less trials mean faster exploration, which relieves the robot of load. The blue bars denote average of flying distance of a ball, i.e., throwing performance. There are three



(a) One-dimensional search space



(b) Two-dimensional search space

Fig. 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.

results in the proposed optimization through synchronization and desynchronization: shoulder (left), waist (middle) or knee (right) was differentiated from other DoFs in the second stage.

It is noted that the results of both proposed methods show less trials than the existing methods. We can find that all results of Hill Climbing show high throwing performance (see Fig.6(a)), which are equivalent to global optimum: the dashed lines (59cm) in Fig. 6. The results of proposed PSO through synchronization are less variance and nearly the same performance as existing PSO with $N = 8$. Therefore, the proposed method can reduce the number of trials while maintaining the high performance.

The number of trials in PSO, compared to the results of the proposed Hill Climbing, is much less. However, the throwing performance of PSO with $N = 8$ is worse than global optimum. More particles (e.g., $N = 20$) are required for the same level of throwing performance as Hill Clibming. There esists a tradeoff between performance and number of particles in PSO.

5.2 Objective function

In Fig. 7, we show the objective function obtained by exhaustive search to discuss above result. Fig. 7(a) illustrates one-dimensional objective function, where the optimal t_{init} is explored in the first stage. There are two local maxima: $t_{init} = -1$ and -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 one-dimensional space should be -1 so that the robot can find the $\mathbf{t}^{opt}$ finally. In Hill Climbing, the optimal t_{init} is -1 if initial value is more than -3. On the other hand, a particle swarm found -5 as global best and then all particles move toward -5 in PSO. This is why the optimization by PSO with synchronization of DoFs was worse than by Hill Climbing (see Fig. 6).

The synchronization of the DoFs of the shoulder and the waist simplifies to reach $\mathbf{t}^{opt}$ because the $\mathbf{t}^{opt}$ is $t_s = t_w = -1$. Thus, the Hill Climbing through synchronizing the DoFs of t_s and t_w in the second stage results in the least number of trials as shown the knee's pink bar in Fig. 6(a). The adequate order of releasing the DoFs may be task-dependent.

6 Discussion

6.1 Necessity of optimization in real world

It is hard for a robot to acquire a skilled throwing. There exists a gap between the real and the virtual world even if we use a realistic simulator or make a dynamical mathematical model of a robot. One of the differences originates from the environmental complexity. The robot's body, for example, interacts with the ball during 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 slight delay makes a difference of performance. After all, it is necessary for acquisition of skilled behavior to optimize the parameters in high-dimensional space using a real robot.

6.2 Synchronization and desynchronization in human skilled behaviors

We demonstrated that optimization of the robot's throwing skill was accelerated by synchronization and desynchronization of the DoFs. The humanoid robot, consequently, could acquire the skilled throwing with less trials (see Fig. 6). The optimal throwing had asynchrony with small differences between DoFs' timing. This asynchrony of DoFs is also observed in human throwing. 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]. In addition, skillful cyclical movements such as clay kneading [11] and samba dance [12] have the slight phase differences between body parts. These studies [11, 12] also showed that there is a process from synchronization to desynchronization of the body parts in acquisition of these periodic skills. The results reported here may suggest that the process in human motor learning has a role of reduction of the motor dimensionality and then accelerates optimization of the movement. Our study, however, does not explain how human optimizes their skills through the process. The desynchronization in human skilled behaviors may result from dynamic interaction between body parts with compliance and environment. More detailed modeling of human motor learning is necessary to expand our approach.

6.3 Possibility of application to other skills

The proposed optimization method were evaluated in throwing task as a case study in this paper. The optimal parameters for throwing in current experiment were $(t_s, t_w, t_k) = (-1, -1, 1)$, which implies that desynchronization with small differences was important for skilled throwing. Athletic behaviors are instantaneous movements, which can be regarded as synchronization of body parts. From a micro perspective, however, a little desynchronization of movements of each body part is required for skilled athletic behaviors (e.g. [14]). In other words, the timing optimized in synchronization of the DoFs is close to the global optimum. Thus, the local maximum reached in the first stage of exploration could be a reasonable initial estimate even if the space dimensionality increases.

We can apply the proposed method to other athletic behaviors if the tasks' optimal parameters exist around synchronized parameters. A slight differentiation of the timing of leg's DoFs 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 applying the proposed method. In addition, velocity of body parts is also important for skilled behavior. We will address the skill acquisition with more parameters such as velocity or acceleration.

7 Conclusion

In this paper, we presented a practical optimization method through synchronization and desynchronization of a robot's body parts. All of the DoFs related to the skill were synchronized in the first stage of learning. Thus, the robot optimized the timing of the start of releasing the ball in one-dimensional space. The DoFs were then desynchronized 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 synchronization of the DoFs resulted in as high performance as the result of optimizing without synchronization 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 synchronized behaviors. Thus the optimization of synchronized DoFs 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 JEAP, which is an acronym for JST ERATO Asada Project, 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. G.S. 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.
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*. Pergamon Press, 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, H.T.A 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.