Biped Gait Optimization Using Estimation of Distribution Algorithm

Lingyun Hu[1,2], Changjiu Zhou[2], Zengqi Sun[1]

[1]*Department of Computer Science and Technology, Tsinghua University, Beijing, 100084, P.R. China*

[2]*School of Electrical and Electronic Engineering, Singapore Polytechnic, 500 Dover Road, 139651, Singapore*

huly02@mails.tsinghua.edu.cn, ZhouCJ@sp.edu.sg, szq-dcs@mail.tsinghua.edu.cn

Abstract - This paper proposes a new biped gait optimization method based on Estimation of Distribution Algorithm (EDA). It is able to explicitly extract global statistical information from the selected solutions and build a posterior probability distribution model of promising solutions based on the extracted information. Biped gait for a nine-link robot is firstly formulated as a multi-objective optimization problem with consideration of multi-constraints including balance and torque. Optimization parameters are angles at transition poses. Instead of searching the joint space directly, EDA is applied to estimate the probability distribution of each joint degree. By this means, inherent mapping relationship between joint coordinates and cost function can be described in term of probability density. Compared to common intelligent learning method, the proposed optimization method can formulate a proper and feasible combination of impulses by tuning less parameters and visiting less states. The effectiveness of the proposed EDA based biped gait optimization method has been tested on a soccer-playing humanoid robot named Robo-Erectus. Experiment results demonstrate that the learned trajectory makes a good balance between stability and energy cost in short learning epochs.

Index Terms - Biped robot, gait optimization, Estimation Distribution Algorithm, probability distribution.

I. INTRODUCTION

Current research on humanoid robots and biped locomotion is one of the most exciting topics in the filed of robotics. As a basic research topic, stable and intelligent biped trajectory generation attracts one's interests from the beginning. Since 1985, Kawarmura *et al*. had proposed a learning method to precisely follow the desired trajectory of each joint [1]. Later in 1992, Li *et al*. described an algorithm on using trunk motion to compensate the lower-limbs motion and to track the desired Zero Moment Point trajectory as close as possible [2]. Following this direction, many studies [3]-[5] employed such idea of deriving the hip motion to achieve the predefined desired ZMP trajectory. However, the actual walking stability of these methods depends greatly on the following three aspects: 1) Reasonability of the desired ZMP trajectory；2) Precision of the mathematical model; 3) Span and variation velocity of joint torques.

For a given robot, the third factor is mostly determined by hardware design. Recently, researchers have taken more attention to the first two directions [6]-[9]. A naturally good idea to solve these problems is to equip the robot controller with some learning capabilities, which makes it possible for robots to follow the pre-designed trajectory. Various kinds of intelligent methods like neural networks, fuzzy logic, evolution algorithms and reinforcement learning have been introduced into trajectory learning. For example fuzzy logic is applied at the local control level for tuning local PID gains in [6]. The research showed that the aggregation-decomposition method for stability analysis of overall biped robot was applicable when local subsystems were stabilized with fuzzy regulator. Traditional feedforward neural network was developed as the controller by Juang [7]. Through learning, it could generate trajectory along a pre-defined path. A typical application of evolutionary algorithm in humanoid robotics was presented in [8], where Genetic Algorithm (GA) easily handled the constraints by using the penalty function vector and transformed a constrained problem into an unconstrained one. Zhou *et al*. formulated a reinforcement method for learning the parameters of rhythmic walking to generate a purposive motion [9].

However, the behaviour of these heuristic computation algorithms depends greatly on a large number of associated parameters. For the example of GAs, quality of evolution process is related to the parameters like operators and probabilities of crossing and mutation, population size, the number of generations and so on. Experiences are required to choose the suitable values for fast resolution. Furthermore, selection of best values for these parameters has been suggested to constitute an additional optimization problem. Moreover, sometimes the algorithms may behave worse than simpler search algorithms for sucking in local minima.

A population-based search algorithm namely Estimation of Distribution Algorithms (EDAs) [10] is employed for biped gait generation and optimization in this paper. EDAs try to predict the population movements in the search space as well as to avoid needing too many parameters. The first explicit description of a solution process consisting with suitable updating on probability distributions was given by Muhlenbein and PaaB [11]. Though the study of EDAs is relatively new, there already exist lots of different types of EDAs like Population-Based Incremental Learning (PBIL) [12], Estimation of Multivariate Normal Algorithm (EMNA) [13], Mutual Information Maximization for Input Clustering (MIMIC) [11] and so on. Since PBIL performs successfully on numerous benchmark and real world problems when compared to a variety of standard GAs and hill-climbing

algorithms. It is adopted in this paper to optimize the generated gaits under energy and stability criteria.

The rest of the paper is organized as follows. In the following section, biped pattern and the mechanical models are briefly introduced. Constraints used in gait generation are presented in the section III. In section IV, gait optimization is abstracted as a multi-objective optimization problem with geometric and state constraints. The EDA used to solve the problem is then demonstrated in section V. This is followed by the algorithm simulation results demonstration in section VI. Section VII is concluding remarks.

II. PROBLEM DEFINITION

A. Biped Gait Pattern

This section introduces the gait model of an N_l-link planar biped robot consisting of a torso, two hips and two identical legs with ankles and knees. Biped walking is assumed to limit in sagittal plane and on a level surface. The main phases of one walking cycle are assumed to be swing phase and double support phase as shown in Fig. 1.
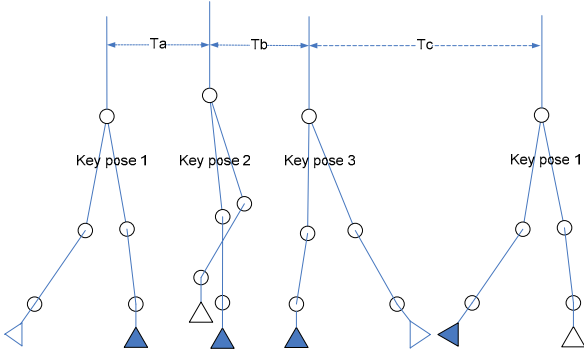

Fig. 1. Gait pattern of one cycle.

- Swing Phase

The swing phase starts from the swing foot toeing off the ground and ends when the heel touches the ground. The swing foot reaches the highest point at T_a and after T_b lands on the ground. The transition moment between swing phase and double support phase is labeled as T_a+T_b.

This phase is modeled as a continuous forward motion in which, the support leg plays as a pivot and the other leg swings forward. Ground friction is assumed to be large enough to ensure no slippage between the supporting foot and walking surface. The mechanism dynamic model of biped robots can be expressed by similar dynamic Newton-Euler equation like manipulators as (1).

$$\hat{M}(q)\ddot{q} + \hat{C}(q,\dot{q}) + \hat{G}(q) = \hat{\tau} \qquad (1)$$

Where $\hat{M}(q) = \{p_{ij}\cos(q_j - q_i)\}_{N_l \times N_l}$ is the square matrix describing full inertia, $\hat{C}(q) = \{-p_{ij}\sin(q_j - q_i)\}_{N_l \times N_l}\dot{q}$ and $\hat{G}(q) = \text{diag}\{-f_i \sin q_i\}_{N_l \times 1}$ are the vector of centrifugal, Coriolis and gravitational moments acting at N_l mechanism

joints respectively, $\hat{\tau}$ is the vector of driving moments at each joint. $J^T(q,\ddot{q})_{6 \times N_l}$ is the corresponding Jacobi matrix of the system. Parameters r_{ij}, s_{ij} and h_i are constants derived by Lagrange's equation. q is the coordinate sets describing the link locomotion in joint space with respect to a world reference frame, $\dot{q}$ is angular velocities. Inequality geometric constraints as shown in the next section are used to avoid collisions between ground and the swing foot.

- Double Support Phase

The double support phase begins at the moment that the front heel touching the ground and ends when the rear toe sways off at $T_a+T_b+T_c$ (as shown in Fig. 1). End configuration of this phase initiates the swing phase of the next gait cycle.

Landing impact happens at the beginning of double support phase. Transition between the two phases is assumed to take place in an infinitesimal length of time. This assumption entails the use of a rigid model to describe the impact between the swing leg and the ground [14]. The impulsive forces may cause an instantaneous velocity change in the generalized coordinates while the positions remain continuous. It can be described as (2).

$$\hat{M}(q)\ddot{q} + \hat{C}(q,\dot{q}) + \hat{G}(q) = \hat{\tau} + \partial F_e \qquad (2)$$

Where F_e is the integral of contact impulse at the impact instant. Since the stance leg is assumed to lift from the ground without interaction, only the external forces at the swing foot F_e need to be considered. By supposing that the actuators are not impulsion, the velocity right after the impact $\dot{q}^+$ can be represented as $\dot{q}^+ = \Delta(\dot{q}^-)$, where $\dot{q}^-$ is the velocity right before the impact and Δ is a continuous function with unique solution.

After the landing impact moment, the mechanical kinematic chain in double support phase can be considered as a manipulator with end position constraints. The dynamical model is given in (3).

$$\hat{M}(q)\ddot{q} + \hat{C}(q,\dot{q}) + \hat{G}(q) = \hat{\tau} + J^T(q,\ddot{q})F \qquad (3)$$

Where F is the vector of external forces and moments acting at the landing foot of the robot.

Totally three key poses are chosen in one cycle. Phases between these key poses are approximated by third order spline functions. A kind of these types of techniques has been implemented using splines of class C^2 in [15]. Such strategy guarantees the second order derivatives at every point, especially at connection points. Consequently, actuating torques is continuous at knots.

B. Mathematical Model

The mathematical model of the biped locomotion proposed in [14] is modified in this paper. The spatial model defined in internal coordinate state space as shown in Fig. 2 will be used to synthesize the dynamic control of the biped mechanism and to verify the simulation research result. Definition of angular coordinates and disposition of the masses are also indicated in Fig. 2.
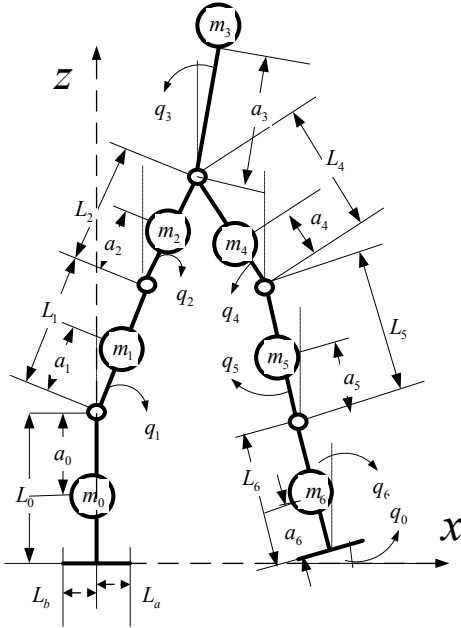
Fig. 2. Mechanical model of the robot.

III. CONSTRAINTS

Constraints are divided into two categories, namely geometric and state constraints in this section according to the definition space.

A. Geometric Constraints

Geometric constraints (GCs) ensure the feasibility of biped gaits from the point of physical structure. Since one gait can be regarded as a pose sequence that is generated by interpolating key poses, GCs will be constructed only at key poses. Two geometric constraints are taken into account.

GC1) Positions limitations of knots as (4).

$$A_p \leqslant p \leqslant B_p \qquad (4)$$

Where $p=[p_i(t)]^T$, $A_p = [A_{p_i}(t)]^T$, $B_p = [B_{p_i}(t)]^T \cdot p_i(t)=[x_i(t),$ $y_i(t), z_i(t)]$ denotes center position of the ith link at t moment. $i=1,2,\dots,N_l$. $A_{p_i}(t)$ and $B_{p_i}(t)$ are lower limit and upper limit of $p_i(t)$. For the example of the position of the swing foot p_6 at T_a. $Ap_6(T_a)=[-(L_0+L_1+L_2),\quad -L_h,\quad L_0-a_0],$ $Bp_6(T_a)=[L_0+L_1+L_2,0,D_h]$. L_h is the width of hip, D_h is the max height of swing foot.

GC2) Structure limitations of joint angles as (5).

$$A_q \leq q \leq B_q \qquad (5)$$

Where $q(t)=[q_i(t)]^T$ stands for the joint angle at time t. $A_q = [A_{q_i}(t)]^T$, $B_q = [B_{q_i}(t)]^T$, $i=0,\dots,N_l-1$. $A_{q_i}(t)$ and $B_{q_i}(t)$ are lower limit and upper limit of $q_i(t)$.

B. State Constraints

Three more kinds of state constraints including force, velocity and stability constraints are also taken into consideration in this paper. Details of each constraint are described as follows.

- **Force constraints**

During double support phase, the leg force must satisfy the force constraint **FC1** as shown in (6).

$$\sum_{i=1}^{N_L} m_i p_i = f_R + f_L + \sum_{i=1}^{N_L} m_i g \qquad (6)$$

Where f_R and f_L are foot force at right and left legs respectively. m_i, p_i are the mass and the acceleration of the ith link. Since only the sum $f_R + f_L$ is known during the double support phase, another force constraint **FC2** is taken into assumption as the internal force f_d in the closed loop structure must be minimized. It can be expressed as (7).

$$f_d = min\{F(f_R, f_L)\} \qquad (7)$$

Where F is the function to calculate the internal force.

- **Velocity constraints**

Velocity constraints (**VC**) are considered to guarantee motion smoothness with respect to mechanical limitations of the biped system. VC can be simply written as:

$$A_{qv} \leqslant \dot{q} \leqslant B_{qv} \qquad (8)$$

where $\dot{q} =[\dot{q}_i(t)]^T$, $\dot{q}_i(t)$ denotes velocity of the ith actuated joint, $A_{qv} = [A_{qv_i}(t)]^T$ and $B_{qv} = [B_{qv_i}(t)]^T$ are lower and upper boundaries of the velocity $\dot{q}_i(t)$.

- **Stability constraint**

To achieve a dynamically stable gait, ZMP is usually used as a criterion [1-5]. The ZMP is defined as the point on the ground at which the net moment of the inertial forces and the gravity forces has no component along the horizontal axes [16]. It can be represented as (9) and (10).

$$x_{zmp} = \frac{\sum\{m_i(\ddot{z}_i + g)x_i - m_i\ddot{x}_i z_i - (I_i\ddot{\theta}_i)_y\}}{\sum m_i(\ddot{z}_i + g)} \qquad (9)$$

$$y_{zmp} = \frac{\sum m_i(\ddot{z}_i + g)y_i - m_i\ddot{y}_i z_i + (I_i\ddot{\theta}_i)_x\}}{\sum m_i(\ddot{z}_i + g)} \qquad (10)$$

So the ZMP constraints (**ZC**) should satisfy:

$$A_{zmp} \leqslant P_{zmp} \leqslant B_{zmp} \qquad (11)$$

Where $p_{zmp}=[p^i_{zmp}(t)]^T$, $p^i_{zmp}(t)=(x^i_{zmp}(t),y^i_{zmp}(t),\ 0)$ is position of the ith ZMP, $A_{zmp} = [A_{zmp_i}(t)]^T$ and $B_{zmp} = [B_{zmp_i}(t)]^T$ are lower and upper boundaries of the stable region respectively.

IV. PERFORMANCE CRITERION

The proposed constrained dynamical model with geometric and state constraints provides the searching space boundary. Based on the proposed framework, an optimization procedure will be developed to select a continuous feasible solution by minimizing an appropriate cost function. Since analytical property of the performance criterion affects the optimal searching in terms of either dynamics or stabilities. ZMP displacement Υ_1 and Energy cost Υ_2 are chosen to form the cost function. Performance criterion is defined as:

Minimize $\quad \Upsilon = \Upsilon_1 + \Upsilon_2$,

Subject to
$\quad$ **GC1** $\quad A_p \leq P \leq B_p$
$\quad$ **GC2** $\quad A_q \leq q \leq B_q$

FC1 $\sum_{i=1}^{N_L} m_i p_i = f_R + f_L + \sum_{i=1}^{N_L} m_i g$

FC2 $f_d = min\{F(f_R, f_L)\}$

VC $A_{qv} \leq \dot{q} \leq B_{qv}$

ZC $A_{zmp} \leq P_{zmp} \leq B_{zmp}$

Where $\Upsilon_1 = \int_{t=0}^{Ta+Tb+Tc} \left\| p_{zmp}(t) - p_{zmp}^d(t) \right\|^2 dt$ estimates the Euclidean distance between the actual ZMP trajectory p_{zmp} and the desired one p_{zmp}^d. $\Upsilon_2 = \int_{t=0}^{Ta+Tb+Tc} \tau(t) dt$ represents the dynamic load during one gait cycle $[0, T_a+T_b+T_c]$.

Let s_n be the total number of discrete sample poses in one gait. Thus the cost function can be defined as (12).

$$C(\bullet) = \beta_f f(\bullet) + \beta_g g(\bullet) = \sum_{i=1}^{s_n} (\beta_f \mathbb{N}(f_i(\bullet)) + \beta_g \mathbb{N}(g_i(\bullet))) \quad (12)$$

Where $f_i(\bullet) = \frac{1}{2} \left\| p_{zmp}^i - p_{zmp}^d \right\|^2$ calculates the Euclidean distance between the actual ZMP and the desired ZMP at the ith sample index. $g_i(\bullet) = \sum_{j=1}^{N_q} \tau_{ij}$ summarizes the torque of all actuated joints at the ith sample index. β_f and β_g are weights satisfying $\beta_f + \beta_g = 1$ to balance between stability and energy cost. $\mathbb{N}$ is the normalization operator.

Appendix I proves that ZMP trajectory will satisfy third order spline interpolation function if joint trajectories are generated by the same function. That is, $f_{j=ks_i}(\bullet)$ and $f_{j=ks_{i+1}}(\bullet)$ makes a difference with $f_{ks_i < j < ks_{i+1}}(\bullet)$. So the proposed algorithm improves only key states to minimize the desired cost function.

V. EDA

The proposed algorithm adopts a new evolutionary algorithm EDA [10]. It is based on the probabilistic model where selection and recombination of building blocks of GAs are replaced by generating new solutions through sampling the probability distribution which is calculated from the selected promising solutions. The probabilistic model is in an explicit model of promising regions of the searching space.

As the input coordinates $q_i (i=1,…,N_l)$ of the N_l links in joint space are considered to satisfy multivariate normal distribution $\Pi_{i=1}^{N_l} p_i(q_i)$, which is the product of N_l independent univariate normal distributions $p_i(q_i)$, the population of solution is replaced by two vectors. One is the mean values of Gaussian normal distribution μ_i and the other is the standard deviation σ_i for each optimized variable q_i. No interactions among the variables are covered. After generating a number of new solutions, the mean values μ_i are shifted towards the best of the generated solutions and the standard deviation σ_i is reduced to make future exploration of the search space narrower.

The algorithm can be summarized as follows.

1. Give a multivariate Gaussian distribution of mean μ_i and standard deviation σ_i.
2. While termination condition is not met:
   a) Generate a population of P individuals.
   b) Select the best Q individuals from parent generation k's population.
   c) Update the multivariate Gaussian distribution using the Q selected individuals. The mean and standard deviation of each univariate distribution are updated in the following ways as (13) and (14).

$$\mu_i(k+1) = (1-\alpha)\mu_i(k) + \alpha(\mu_{i,best}(k) + \mu_{i,2ndbest}(k) - \mu_{i,worst}(k)) \quad (13)$$

$$\sigma_i(k+1) = (1-\alpha)\sigma_i(k) + \alpha\sqrt{\frac{1}{Q}\sum_{i=1}^{Q}(\mu_i^j(k) - \tilde{\mu}_i^j(k))^2} \quad (14)$$

Where $\mu_{i,best}(k)$, $\mu_{i,2ndbest}(k)$ and $\mu_{i,worst}(k)$ are the values of the best, second best and worst individual (with respect to the cost function) of the ith variable at iteration k. μ_i^j $j=1,…,Q$ are the Q best individuals for the ith variable and $\tilde{\mu}_i^j$ is their mean. α is the learning rate.

VI. SIMULATION RESULTS

To show the effectiveness of the proposed algorithm, we apply it on the digital robot model of humanoid robot named Robo-Erectus (RE) as shown in Fig.3 [17]. Height and weight of RE are 600 mm and 4.6 kg respectively. Totally 23 DOFs are designed in RE as 6 per leg, 4 per arm and 1 for head, 2 for waist. They are driven by servo motor with maximum torque as $30kg \times cm$. One notebook powered with window XP operation system is set on RE for online control. It is one of the foremost leading soccer-playing humanoid robots, in the RoboCup Humanoid League. (www.robo-erectus.org)

Fig. 3 Humanoid robot Robo-Erectus.

Based on the robot, a simplified model as shown in Fig. 2 is constructed for simulation experiment. Basic structure parameters of the model are N_l=7, N_{kp}=4, s_n=10, L_h=10cm, L_0=4cm, L_1= L_2=8cm, L_3=20cm, M_0=0.1kg, M_1= M_2=0.2kg, M_3=1kg. Desired ZMP is set as the middle line in the stable region. Other parameters used in EDA are designed as Q=40, P=20, α=0.01. EDA stops learning after 200 iterations.

Fig. 4 exhibits the probability distribution $Pro(q_1)$ of q_1 in the ten looping times. Probability distribution functions $Pro(q_i)$ of other coordinate parameters q_i converge similarly like $Pro(q_1)$. i=1,…,N_l. Variation of torque and ZMP in the learning procedure are shown in Fig. 5 and Fig.6 respectively, from which it can be found that both of them converge to reasonable values after EDA learning.
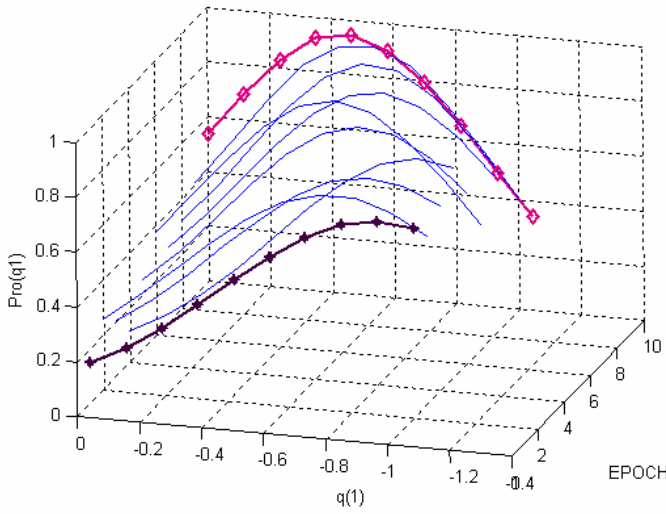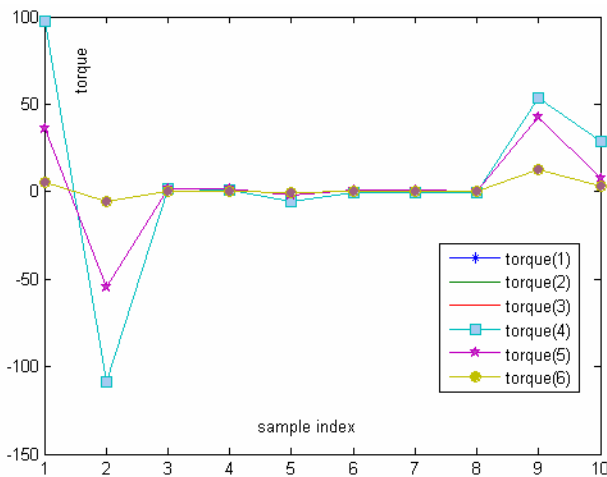

Fig. 4 Probability model of q(1).
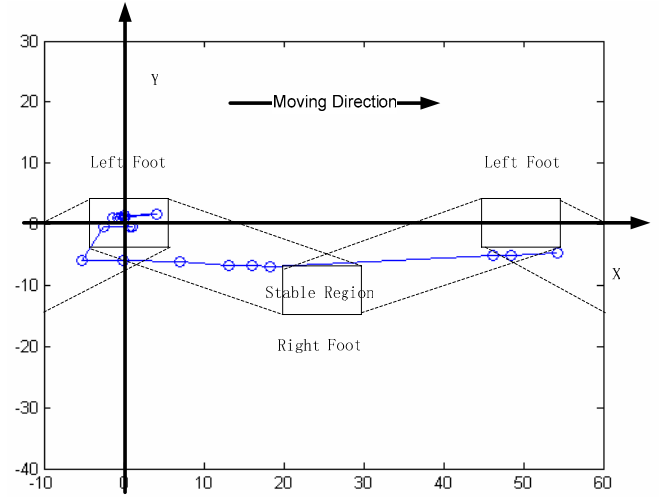

Fig. 5 Torque of learned gait in one cycle.

Fig.6 ZMP trajectory of learned gait in one cycle.

VI. CONCLUSIONS

We propose a general biped gait optimization algorithm based on EDA in this paper. Gait synthesis of the biped locomotion is formulated as a multi-objective optimization problem. By supposing that the probability distribution of each joint degree satisfies Gaussian distribution function, EDA is applied to find the stable and low energy cost joint degree permutation. To the best of our knowledge, the proposed algorithm is the first biped gait optimization method using EDA. It can be easily extended on more joints for more complex robot structure, which will be discussed in the future. We will also look at how to use computational learning approaches, e.g. fuzzy reinforcement learning [19] to further optimize biped gaits.

APPENDIX I

In this section, we will prove that if joint trajectory (x_j^i, y_j^i, z_j^i) is generated by third order spline interpolation functions (Sx_j^i, Sy_j^i, Sz_j^i), the corresponding ZMP trajectory (x_{zmp}^j, y_{zmp}^j) also satisfies third order spline interpolation functions (Sx_{zmp}^j, Sy_{zmp}^j).

Proof: Take the example of x_{zmp}^j.

Since $z_j^i(t) = S_{z_j^i}(t) = A_{z_j^i}t^3 + B_{z_j^i}t^2 + C_{z_j^i}t + D_{z_j^i}$

$$x_j^i(t) = S_{x_j^i}(t) = A_{x_j^i}t^3 + B_{x_j^i}t^2 + C_{x_j^i}t + D_{x_j^i}$$

Then $\ddot{z}_j^i(t) = \ddot{S}_{z_j^i}(t) = 6A_{z_j^i}t + 2B_{z_j^i}$

$$\ddot{x}_j^i(t) = \ddot{S}_{x_j^i}(t) = 6A_{x_j^i}t + 2B_{x_j^i}$$

$A_{z_j^i}, B_{z_j^i}, C_{z_j^i}, D_{z_j^i}, A_{x_j^i}, B_{x_j^i}, C_{x_j^i}, D_{x_j^i}$ are parameters produced by the key states and predetermined gait parameters like gait length and gait height. It leads to

$$x_{zmp}^j = \frac{\sum_{i=1}^{n} m_i(\ddot{z}_j^i + g)x_j^i - \sum_{i=1}^{n} m_i\ddot{x}_j^i z_j^i}{\sum_{i=1}^{n} m_i(\ddot{z}_j^i + g)}$$

$$= \frac{\sum_{i=1}^{n} m_i(\ddot{S}_{z_j^i}(t) + g)(S_{x_j^i}(t)) - \sum_{i=1}^{n} m_i\ddot{S}_{x_j^i}(t)(S_{z_j^i}(t))}{\sum_{i=1}^{n} m_i(\ddot{S}_{z_j^i}(t) + g)}$$

$$= \frac{\sum_{i=1}^{n} m_i(6A_{z_j^i}t + 2B_{z_j^i} + g)(A_{x_j^i}t^3 + B_{x_j^i}t^2 + C_{x_j^i}t + D_{x_j^i})}{\sum_{i=1}^{n} m_i(6A_{z_j^i}t + 2B_{z_j^i} + g)}$$

$$- \frac{\sum_{i=1}^{n} m_i(6A_{x_j^i}t + 2B_{x_j^i})(A_{z_j^i}t^3 + B_{z_j^i}t^2 + C_{z_j^i}t + D_{z_j^i})}{\sum_{i=1}^{n} m_i(6A_{z_j^i}t + 2B_{z_j^i} + g)}$$

$$= A_{x_{zmp}}t^3 + B_{x_{zmp}}t^2 + C_{x_{zmp}}t + D_{x_{zmp}}$$

$$= S_{x_{zmp}^j}(t)$$

The same deduction can be carried on y_{zmp}^j.

ACKNOWLEDGMENTS

REFERENCES

[1] S. Kawamura, T. Kawamura, D. Fijino, F. Miyazaki, S. Arimoto, "Realization of Biped Locomotion by Motion Pattern Learning," *Robotics Society of Japan*, vol.3, no.3, pp.5-15, 1985.

[2] Q. Li, A. Takanishi, I. Kato, "Learning Control of Compensative Trunk Motion for Biped Walking Robot Based on ZMP Stability Criterion," *in Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, 1992, pp. 597-603.

[3] Q. Huang, K. Yokoi, S. Kajita, K. Kaneko, *et. al*, "Planning Walking Patterns for A Biped Robot," *IEEE Trans. on Robotics and Automation*, vol. 17, no. 3, pp. 280-290,2001.

[4] R. Kurazume, T. Hasegawa, K. Yoneda, "The Sway Compensation Trajectory for A Biped Robot," *in Proc. IEEE Int. Conf. Robotics and Automation*, 2003, pp. 925-931.

[5] J. Yamaguchi, A. Takanish, I. Kato, "Development of A Biped Walking Robot Compensating for Three-axis Moment by Trunk Motion," *in Pro. IEEE/RSJ Int. Workshop on Intelligent Robotics and Systems*, 1993, pp. 561-566.

[6] M. Vukobratovic, O. Timcenko, "Stability Analysis of Certain Class of Bipedal Walking Robots with Hybridization of Classical and Fuzzy Control," *in Proc. of the 1st ECPD Interna. Conf. on Advanced Robotics and Intelligent Automation*, Athens, Greece, 1995, pp. 290-295.

[7] J. G. Juang, "Minimal Energy Control On Trajectory Generation", *in Pro. Interna. Conf. on information intelligence and systems*, 1999, pp. 204-210.

[8] G. Capi, S. Kaneko, K. Mitobe, L. Barolli, Y. Nasu, "Optimal Trajectory Generation for A Prismatic Joint Biped Robot Using Genetic Algorithms," *Robotics and Autonomous Systems*, 38, pp.119-128, 2002.

[9] C. Zhou, Q. Meng, "Reinforcement Learning with Fuzzy Evaluative Feedback for A Biped Robot," *in Proc. IEEE Int. Conf. Robotics and Automation*, 2000, pp. 3829-3834.

[10] P. Larranaga, J. A. Lozano, *Estimation of Distribution Algorithms. A New Tool for Evolutionary Computation*, Kluwer Academic Publishers, 2001.

[11] H. Muhlenbein, G. PaaB, "From recombination of genes to the estimation of distributions I. binary parameters," *Lecture Notes in Computer science 1141: Parallel Problem Solving from Nature-PPSN IV*, Springer Verlag, Berlin, pp. 178-187, 1996.

[12] J.R. Greene, "Population-based Incremental Learning as A Simple Versatile Tool for Engineering Optimization," *in Proc. of the first Int. Conf. on EC and Applications*, 1996, pp. 258-269.

[13] P. Laggranaga, J.A. Lozano, E. Bengoetxea, "Estimation of Distribution Algorithms Based on Multivariate Normal and Gaussian Network," *Technical Report KZZA-IK-1-01, Department of Computer Science and Artificial Intelligence*, University of the Basque Country, 2001.

[14] J.W. Grizzle, G. Abba, F. Plestan, "Asymptotically Stable Walking for Biped Robots: Analysis Via Systems with Impulse Effects," *IEEE Trans. On Automatic Control*, vol. 46, No. 1, pp. 51-58, 2001.

[15] C. Chevallereau, Y. Aoustin, "Optimal Reference Trajectories for Walking and Running of A Biped Robot," *Robotica*, vol. 19, pp. 557-569, 2001.

[16] M. Vukobratovic, B. Borovac, D. Surla and D. Stokic, Biped Locomotion: Dynamics Stability, Control, and Application, Springer-Verlag, New York, 1990.

[17] C. Zhou, P.K. Yue, "Robo-Erectus: A Low Cost Autonomous Humanoid Soccer Robot," *Advanced Robotics*, vol.18, no.7, pp. 717-720, 2004.

[18] L. Hu, C. Zhou and Z. Sun, "Optimizing biped gait using probability model," in *Proc. Int. Conf. on Computational Intelligence, Robotics and Autonomous Systems* (CIRAS2005), 2005 (Accepted).

[19] C. Zhou, "Robot learning with GA-based fuzzy reinforcement learning agents," *Information Sciences*, vol. 145, pp.45-68, 2002.