

Active Balance Control of Humanoid Locomotion Based on Foot Position Compensation

Chengju Liu¹, Tong Zhang¹, Ming Liu², Qijun Chen^{1*}

1. School of Electronics and Information Engineering, Tongji University, Shanghai 201804, China

2. School of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Hong Kong 999077, China

Abstract

A foot positioning compensator is developed in this paper for a full-body humanoid to retrieve its balance during continuous walking. An online Foot Position Compensator (FPC) is designed to improve the robustness of biped walking, which can modify predefined step position and step duration online with sensory feedback. Foot placement parameters are learned by the FPC based on the Policy Gradient Reinforcement Learning (PGRL) method. Moreover, the FPC assists the humanoid robot in rejecting external disturbances and recovering the walking position by re-planning the trajectories of walking pattern and the Center of Mass (CoM). An upper body pose control strategy is also presented to further enhance the performance of humanoid robots to overcome strong external disturbances. The advantages of this proposed method are that it neither requires prior information about the walking terrain conditions, nor relies on range sensor information for surface topology measurement. The effectiveness of the proposed method is verified *via* Webots simulation and real experiments on a full-body humanoid NAO robot.

Keywords: humanoid walking, active balance, Foot Positioning Compensation (FPC), Policy Gradient Reinforcement Learning (PGRL)

Copyright © Jilin University 2020.

1 Introduction

Legged robots have an anthropomorphic appearance and more flexible mobility compared with wheeled robots. When exposed to unknown external disturbances, such as irregular terrain or pushes, the ability to perform robust and adaptive walking is the most basic and important precondition to fulfill various tasks^[1–3]. Biped locomotion is generally achieved by designing the trajectory of the Center of Mass (CoM) so that the actual Zero Moment Point (ZMP) follows the desired one^[4,5], which is decided by a pre-defined foot position. However, this planning approach is executed offline typically, and the walking gait can't be adjusted flexibly. In online implementations, the robot is typically represented by simplified dynamics models, such as Cart-table model and Linear Inverted Pendulum Model (LIPM)^[6–11], resulting in larger approximation errors than that of encountered in the actual humanoid robot dynamics.

To guarantee the adaptive walking in unknown conditions in the presence of external disturbances, such as pushes or terrain irregularities, a great number of

control strategies have been proposed to regulate the CoM or tracks ZMP reference trajectories through robust feedback schemes^[12–17]. Castano *et al.*^[15] proposed a locomotion control strategy using feedback information of the actual CoM state to compensate the foot placements through orbital energy based on the LIPM to recover the motion velocity. Hu *et al.*^[16] presented a real time learning framework to improve the robustness of ZMP-based biped walking controllers, which using iterative learning control to learn a feed forward compensative ZMP from measured ZMP errors during repetitive locomotion. Mason *et al.*^[17] extended simplified walking pattern generators to exploit hand contacts when stability cannot be maintained by stepping alone. Biological inspired method, such as Central Pattern Generator (CPG) methods are also used to online regulate the CoM or ZMP reference trajectories to improve the adaptive and robustness of biped walking^[18–22]. In our previous work, an online CoM trajectory modulation method based on CPGs is proposed using sensory feedback to real slop terrain walking^[19]. In Ref. [20], in order to improve the adaptability, foot trajectory and

*Corresponding author: Qijun Chen, Senior member, IEEE
E-mail: qjchen@tongji.edu.cn

CoM trajectory generators are proposed to realize uneven terrain adaptive walking. These approaches imitate the instantaneous reactive balance strategies of human beings.

Recently, increasingly attention has been devoted to strong disturbance rejection in humanoid robot locomotion with adjustable foot positioning. The adjustment of foot positioning changes the force region available for the CoM control. Therefore, appropriate foot positioning may provide corresponding forces to regulate the CoM trajectory as desired. To reject external disturbances during continuous locomotion, Wieber *et al.* presented an approach that can generate online walking with adjustable step duration and foot positioning using the linear model predictive-control strategy^[23–25]. Nishiwaki *et al.* proposed the methods for humanoid locomotion in which balance was maintained by adjusting the duration of the current step, next step position, and the reference ZMP trajectory^[26]. The effectiveness of the approach was validated by using a real humanoid robot in several experiments. Feng *et al.* presented an online foot step optimization strategy to guarantee stable walking of a robot subject to external disturbances by model predictive control^[27]. This strategy was implemented and verified on the Atlas robot. Kryczka *et al.*^[28] proposed a gait pattern regeneration method that can control step positions and step timing to recover from disturbances. Learning-based techniques are typically utilized in amending the step placements to maintain balance. Missura *et al.*^[29,30] proposed an online learning method to balance a humanoid robot against strong interferences and follow a desired step size. Based on the adjustable non-parametric regression model, we proposed a heteroscedastic sparse GP (HSpGP) based FPC approach to improve the robustness of biped walking^[31,32]. The proposed HSpGP-based FPC approach is nonparametric, which need not to precise modeling the dynamic relationship between the humanoid robot state and the modified steps.

This study focuses on enhancing the locomotion stability and robustness of a humanoid robot exposed to strong external disturbances by compensating the foot position in real time. The main approach is that FPC parameters are learned from torso angle errors and measured CoM through sensory feedback information

during every locomotion cycle. The approach optimizes the analytic solutions to step positioning by using the PGRL strategy based on the Cart-table model^[33,34]. Thus, a novel method is proposed to modify foot placement in real time to regain balance against disturbances. The requirements on the control system are given as follows:

(1) The dynamics of a humanoid robot are approximated as a cart on a massless table. The legs of the humanoid robot are assumed to be a weightless scalable limb of the table in the cart-table model, and we will not consider its kinematic constraints. If a series of next foot positions is given, a reference ZMP can be generated to meet the requirement of dynamic stability and based on the Cart-table model the required CoM trajectory can be obtained online using preview control.

(2) To improve accuracy in tracking the CoM, preview control is used to generate walking patterns, that is, a preview controller with an observer is adopted as foundation in our work.

(3) Humanoid walking quality is evaluated by determining the capacity to recover balance when subjected to a strong external disturbance. Although an optimal footstep placement ensures the dynamic stability of the robot, it may tip over while walking in an uncertain environment, *e.g.*, uneven terrain or slopes, as the actual body posture does not conform to the desired posture due to tracking errors. Therefore, a novel strategy is formed to compensate the upper body posture by controlling the inclination of the robot torso online with hip joints.

The remainder of the paper is structured as follows. The dynamic state of the humanoid robot is presented in section 2. The applied concept of the preview controller and observer are presented in section 3. The design approach to the FPC is presented in section 4. The control of the upper body pose using the hip method is developed in section 5. In section 6, the effectiveness of the method is demonstrated by simulation and experiments using the NAO robot. The paper concludes in section 7, which includes an outline of future work.

2 Dynamic state of the humanoid robot

The modeling of the dynamic state of the robot in conjunction with the sensor configuration is described in this section.

2.1 Cart-table model and coordinate systems

The Cart-table model is applied to represent the main dynamics of the humanoid robot. The Cart-table model is shown in Fig. 1, consisting of a massless table with a moving cart. When the cart is moving at a proper speed the table can keep its balance. The cart-table model is adopted to simplify the robot dynamics and figure out the CoM trajectory. For the given position of the CoM of the cart and acceleration g , the ZMP can be described as:

$$p_x = q_x - \frac{\ddot{q}_x}{g} z_c, \tag{1}$$

where q_x and \ddot{q}_x represent the position and acceleration of the CoM, respectively, along the x -axis, g is the gravitational acceleration, and $z_c = q_z - p_z$ denotes the height of the CoM, which can be considered constant. q_z and p_z are the position of CoM and ZMP in z -axis direction.

The state of the humanoid robot can be represented as a vector $\mathbf{X} = (q_x, \dot{q}_x, p_x)^\top$. The equation for the y -axis is analogous to that of the x -axis. The control system presented here is common for humanoid robots. In this work, a NAO robot, is used as the platform. The frequency of walking control and sensor data module is 20 Hz. The weight of each leg is almost the same as that of the torso, and the links are made of plastic. Furthermore, the deformation of plastic links could be caused by the disturbance.

2.2 Sensors

The robot is equipped with accelerometers and gyroscope. Attitude of the robot can be estimated using complementary filters with these sensors. The CoM and ZMP positions in world coordinates can then be estimated without external equipment. The touch sensors installed under the soles are applied to compute the ZMP as^[35]:

$$p = \frac{p_L f_L + p_R f_R}{f_L + f_R}, \tag{2}$$

where p_L and p_R are the ZMP position calculated for the left and right feet, respectively; f_L and f_R are the forces of the left and right feet, respectively, and are measured by Force-Sensitive Resistors (FSRs).

The body attitude θ of the NAO robot is used for

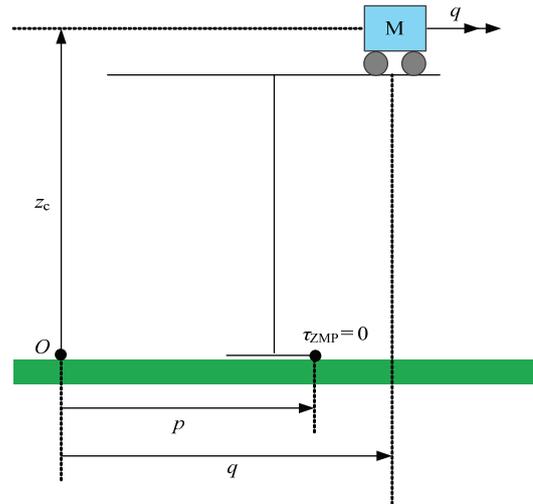


Fig. 1 The Cart-table model.

the posture control of the upper body and is estimated using the robot's gyroscope and accelerometer. Assuming that the robot is stationary relative to the ground, and the acceleration sensor is in an ideal state without system error and noise. The output signal of the triaxial sensor is the vector decomposition of the gravitational acceleration g in the triaxial direction. The current body rotation angle around the x and y axes can be calculated based on the real-time data collected by the accelerometer. The formulas are:

$$\theta_{accx} = a \tan 2 \left(-c_y, \sqrt{c_x^2 + c_z^2} \right), \tag{3}$$

$$\theta_{accy} = a \tan 2 \left(c_x, \sqrt{c_y^2 + c_z^2} \right), \tag{4}$$

where θ_{acc} is the body attitude calculated by the accelerometer. c_x , c_y , and c_z are the output components of the accelerometer in triaxial direction.

When the robot is in motion, the body posture angles in both x and y directions can be calculated by integrating the angular velocity by time. The formulas are:

$$\theta_{gyrox} = G_{yro} X \cdot dt + \text{last_}\theta_{gyrox}, \tag{5}$$

$$\theta_{gyroy} = G_{yro} Y \cdot dt + \text{last_}\theta_{gyroy}, \tag{6}$$

where θ_{gyro} is the body attitude calculated by the gyroscope, dt is the sampling time, $G_{yro} X$ and $G_{yro} Y$ are the angular velocities in the current state and $\text{last_}\theta_{gyrox}$ and $\text{last_}\theta_{gyroy}$ are the accumulated body angles in previous states.

These two methods for calculating the body posture angle are based on ideal conditions. However, due to the high-frequency oscillation of the robot system and the influence of sensor noise, the calculated body posture angle cannot objectively and accurately reflect the posture of the robot. Because the acceleration sensor provides static information, a compensation filter is designed in this paper, and the gyroscope reflects the dynamic characteristics of movement. The calculated body posture angle of the acceleration sensor is filtered using a low-pass filter. The low-frequency data collected by the gyroscope is filtered through a high-pass filter. The final body posture angle is the weighted summation of the two filtered angles. In the following equation, GainL and GainH are weights that are set to 1 in the experiment.

$$\theta = \text{GainL} \times \text{Low Filter}(\theta_{\text{acc}}) + \text{GainH} \times \text{High Filter}(\theta_{\text{gyro}}). \tag{7}$$

3 Preview controller and observer

Preview control is an easily engineered theory to generate an optimal CoM trajectory that follows preplanned ZMP trajectory. This theory is widely used in many complex humanoid robots^[36-41]. According to the theory of preview control, the current control input of the system is connected with not only the current states but also the future desired output of the system. This theory allows the control system to take the future states of the ZMP into consideration to improve the system tracking performance of the CoM trajectory. By preview control, we can predict the tracking performance of the current control input to future output and obtain the optimal control input.

Here, we discuss only *x*-axis motion. The results for the *y*-axis are identical. Let p_x and q_x be the ZMP and CoM positions, respectively, on the *x*-axis. Define u as the speed of the ZMP, which is the output of the system, with a sampling time period Δt , Eq. (1) can be rewritten as a discrete-time system as follows:

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{A}_0 \mathbf{x}_k + \mathbf{b}u_k \\ p_k = \mathbf{C} \mathbf{x}_k \end{cases}, \tag{8}$$

where $\mathbf{x}_k = [q_x(k\Delta t), \dot{q}_x(k\Delta t), p_x(k\Delta t)]^T$, $u_k = u(k\Delta t)$, $p_k = p_x(k\Delta t)$, $\mathbf{b} = [0 \ 0 \ \Delta t]^T$, $\mathbf{C} = [0 \ 0 \ 1]^T$, and

$$\mathbf{A}_0 = \begin{bmatrix} 1 & \Delta t & 0 \\ \frac{g}{z_c} \Delta t & 1 & -\frac{g}{z_c} \Delta t \\ 0 & 0 & 0 \end{bmatrix}.$$

An evaluation function termed the performance index J is defined as:

$$J = \sum_{j=1}^{\infty} \left\{ \mathbf{Q}_e [p^{\text{ref}}(j) - p(j)]^2 + \Delta \mathbf{x}^T(j) \mathbf{Q}_x \Delta \mathbf{x}(j) + \mathbf{R} \Delta u^2(j) \right\}, \tag{9}$$

where \mathbf{Q}_e , \mathbf{Q}_x and \mathbf{R} are positive weighting coefficients, and $p^{\text{ref}}(j)$ denotes the j th step's reference ZMP.

According to preview control theory and using N frames data of a future ZMP reference, the evaluation of function J can be minimized as:

$$\mathbf{u}(k) = -\mathbf{G}_I \sum_{i=1}^k [\mathbf{C} \mathbf{x}(i) - p^{\text{ref}}(i)] - \mathbf{G}_x \mathbf{x}(k) - \sum_{j=1}^N \mathbf{G}_d(j) p^{\text{ref}}(k+j). \tag{10}$$

The gains of the preview controller of Eq. (10) are calculated as follows:

$$\mathbf{G}_I = [\mathbf{R} + \mathbf{B}^T \mathbf{P} \mathbf{B}]^{-1} \mathbf{B}^T \mathbf{P} \mathbf{I}, \tag{11}$$

$$\mathbf{G}_x = [\mathbf{R} + \mathbf{B}^T \mathbf{P} \mathbf{B}]^{-1} \mathbf{B}^T \mathbf{P} \mathbf{F}, \tag{12}$$

$$\mathbf{G}_d(j) = -[\mathbf{R} + \mathbf{B}^T \mathbf{P} \mathbf{B}]^{-1} \mathbf{B}^T [\mathbf{A}_c^T]^j \mathbf{P} \mathbf{I}, \tag{13}$$

$j = 0, 1, \dots, N,$

where $\mathbf{B} = \begin{bmatrix} \mathbf{C} \mathbf{b} \\ \mathbf{b} \end{bmatrix}$, $\mathbf{I} = [1 \ 0 \ 0]^T$, $\mathbf{F} = \begin{bmatrix} \mathbf{C} \mathbf{A}_0 \\ \mathbf{A}_0 \end{bmatrix}$, $\mathbf{A} = [\mathbf{I} \ \mathbf{F}]$,

and \mathbf{P} is the solution of the *Riccati* equation as:

$$\mathbf{P} = \mathbf{A}^T \mathbf{P} \mathbf{A} - \mathbf{A}^T \mathbf{P} \mathbf{B} [\mathbf{R} + \mathbf{B}^T \mathbf{P} \mathbf{B}]^{-1} \mathbf{B}^T \mathbf{P} \mathbf{A} + \mathbf{Q}. \tag{14}$$

where $\mathbf{Q} = \begin{bmatrix} \mathbf{Q}_e & 0 \\ 0 & \mathbf{C}^T \mathbf{Q}_x \mathbf{C} \end{bmatrix}$. For a given ZMP trajectory $p^{\text{ref}}(j)$, the optimal u can be obtained, and the CoM state of the next frame can be generated according to Eq. (8). Thus, the joint commands can be figured out by inverse kinematics. The performance of this preview controller is illustrated in Fig. 2, which shows the CoM and ZMP trajectory based on the preview control method.

The CoM reference trajectory begins to move before

the ZMP reference trajectory changes in every walking step. The comparison of the actual ZMP and reference ZMP trajectory illustrates that the tracking accuracy of the system is generally good. However, the reference CoM trajectory cannot be tracked precisely due to the model error between the simplified Cart-table model and full-body humanoid robot.

The aforementioned preview control is solely an open-loop control method. Additional parameters, such as the time sampling period, the time ratio of the single support phase to the double support phase, and the displacement from the reference ZMP to the origin of the supporting foot, are required in actual humanoid robot control. With a fine-tuned parameter set, the robot can manage to walk on a generally flat plane and can handle small external disturbances. However, if the parameters are not fine-tuned or the robot encounters a relatively large external impulsive disturbance, it will fall easily. Unfortunately, this situation will remain true because of the difference between the simplified LIPM and the complete dynamics of the humanoid. The elasticity of links and other mechanical issues will cause the robot to be increasingly unstable. This problem is typically addressed by an observer.

As the present CoM and ZMP trajectories can be figured out according to feedback information obtained by the sensors, the state vector \hat{x} estimation can be utilized by the observer.

Hence, the system's output is described as:

$$y(k) = C_m \cdot x(k) \hat{=} \begin{bmatrix} q_x^{mea}(k) \\ p_x^{mea}(k) \end{bmatrix}, \quad (15)$$

where $q_x^{mea}(k)$ and $p_x^{mea}(k)$ denote the measured CoM position and ZMP position, respectively, and

$$C_m = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (16)$$

Taking sensor measurement into consideration, the discrete time system with an observer can be modified as:

$$\hat{x}_{k+1} = A_0 \hat{x}_k + L[y_k - C_m \hat{x}_k] + b \hat{u}_k, \quad (17)$$

where L is an observer gain, which can be obtained by the discrete LQ regulator method.

Omnidirectional walking can be easily realized based

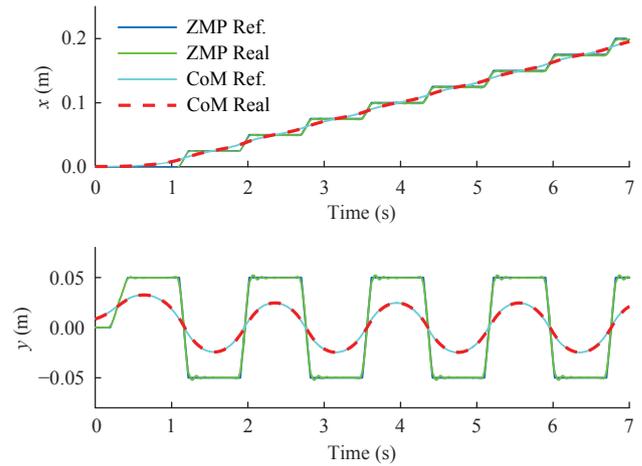


Fig. 2 ZMP and CoM trajectories along the x-axis and y-axis using preview control.

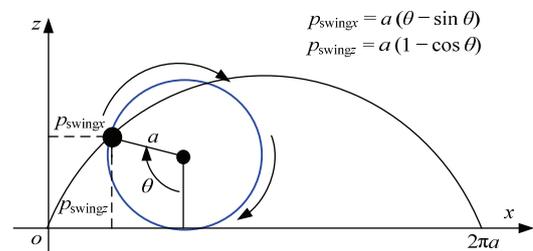


Fig. 3 Cycloidal trajectory diagram.

on this preview control scheme. Fig. 3 shows the method of generating a swing foot trajectory. p_{swingx} and p_{swingy} are the positions of the swing feet in x and z direction, respectively.

Let $p_s = [x_s, y_s, z_s]^T$ and $p_e = [x_e, y_e, z_e]^T$ be the initial and landing positions of the swing foot, respectively. A cycloid curve is adopted for the swing foot trajectory. Given a desired walking speed $v_c = [v_{cx}, v_{cy}]^T$ and turning speed ω , the position on the following time instant can be given as:

$$\begin{cases} p^f = \left[\frac{\alpha}{2} v_{cx}, \alpha v_{cx} + y_s \right], \\ \theta^f = \alpha \omega \end{cases} \quad (18)$$

where $\alpha = T/100$, y_s is the distance between the two feet along the y -axis when the robot stands still, and T represents the locomotion period in seconds. An illustration of omnidirectional walking steps is shown in Fig. 4. As shown, the ZMP trajectory is then generated from the positions of the footsteps.

$$F(\phi) = \frac{1}{n} \sum_{j=1}^n \left(\alpha_x \cdot (|\Delta p_x| + |\hat{x} - x^{ref}|) + \alpha_y \cdot (|\Delta p_y| + |\hat{y} - y^{ref}|) + \beta_x \cdot |\Delta \theta_p| + \beta_y \cdot |\Delta \theta_r| \right), \tag{22}$$

where n denotes the number of interpolations during the single support phase, α_x and α_y represent the weight factors of CoM error along the x - and y -axes, respectively, β_x and β_y denote weight factors of torso angle error along the x - and y -axes, respectively, $\alpha_x + \alpha_y = 1$, and $\beta_x + \beta_y = 1$.

First, the matrix of gain coefficient can be initialized as ϕ^0 by manual tuning or random generation. Next, at each parameter value in ϕ^0 , the partial derivative of $F(\phi)$ can be calculated and p policies can be generated near ϕ^0 by setting ${}_m\phi^0 (m = 1, \dots, p)$:

$${}_m\phi^0 = \phi^0 + [\rho_1, \dots, \rho_k], \tag{23}$$

where $\rho_j (j = 1 \dots k)$ is a disturbance randomly chosen in $\{-\varepsilon_j, 0, \varepsilon_j\}$.

By grouping every ${}_m\phi^0$ into one of three sets as Eq. (19) shown, the matrix of average score can be calculated.

$${}_m\phi^0 \in \begin{cases} G_{-\varepsilon,j} & , \text{ if policy } \rho_j \text{ is chosen by } -\varepsilon_j \\ 0 & , \text{ if policy } \rho_j \text{ is chosen by } 0 \\ G_{+\varepsilon,j} & , \text{ if policy } \rho_j \text{ is chosen by } +\varepsilon_j \end{cases} . \tag{24}$$

Next, the average scores $\bar{F}_{-\varepsilon,j}$, \bar{F}_0 and $\bar{F}_{+\varepsilon,j}$ for the corresponding group, $G_{-\omega_j}$, G_0 and $G_{+\omega_j}$, can be obtained. The adjustment F_n is expressed as:

$$F_n = \begin{cases} 0 & , \text{ if } \bar{F}_{-\varepsilon,j} > \bar{F}_0 \text{ and } \bar{F}_{+\varepsilon,j} > \bar{F}_0 \\ \bar{F}_{+\varepsilon,j} - \bar{F}_{-\varepsilon,j} & , \text{ otherwise} \end{cases} . \tag{25}$$

The approximating gradient value $\nabla \phi^0$ is $\nabla \phi^0 = -F_n / |F_n|$, and the update rule of gain coefficient matrix is $\phi^1 = \phi^0 + \nabla \phi^0 \cdot \eta$, in which the step size factor η is a constant. The previous steps are repeated utilizing an updated matrix of gain coefficient. The termination condition of the learning process is when the number of iterations exceeds a given number N_{iter} . The local optimal solution can be calculated when N_{iter} is large enough. Therefore, the optimal parameters for the FPC can be obtained through the process of learning. The speed of computation can be ensured by executing the learning process during a single support phase.

5 Upper body posture control

Despite the fact that the influence of the upper body posture of the robot on the adjustment of the foot placement is taken into account, the posture of upper body is not controller directly. Therefore, the robot may fall when walking on an uphill slope or terrain. For these reasons, a controller based on upper body posture has been proposed in this work to amend the body posture with the hip joints of robot. Fig. 5 shows the diagram of the presented control method. Refer to Ref. [35] for more details on the motion control process.

In the upper body posture control method, the correction angles of hip joints ($\Delta \theta_{HipPitch}$, $\Delta \theta_{HipRoll}$), based on the PD control method can be presented as:

$$\Delta \theta_{HipPitch} = K_{p1} (\theta_p - \hat{\theta}_p) + K_{d1} (\dot{\theta}_p - \hat{\dot{\theta}}_p), \tag{26}$$

$$\Delta \theta_{HipRoll} = K_{p2} (\theta_r - \hat{\theta}_r) + K_{d2} (\dot{\theta}_r - \hat{\dot{\theta}}_r), \tag{27}$$

where K_p and K_d are the proportional and differential gains of the PD controller, respectively, (θ_p, θ_r) are the ideal angles of body in the x - and y -axes (in this experiments, it was set as $(0, 0)$), respectively, $(\dot{\theta}_p, \dot{\theta}_r)$ are the ideal angular velocities of body, $(\hat{\theta}_p, \hat{\theta}_r)$ indicate the measured body angles, and $(\hat{\dot{\theta}}_p, \hat{\dot{\theta}}_r)$ are the measured angular velocities of body. The control strategy is performed during each moment of the walking period.

6 Experiments

According to the overall control diagram presented in Fig. 5, some simulations are carried out on the Webots platform using NAO robot to prove the validness of the proposed method. The process of learning and some supported experiment are conducted on the Webots. The information required to estimate CoM and ZMP states of robot is attained by accelerometers with additive noise and a simulated gyroscope. The videos of the simulation and experiments can be found in the supplementary materials.

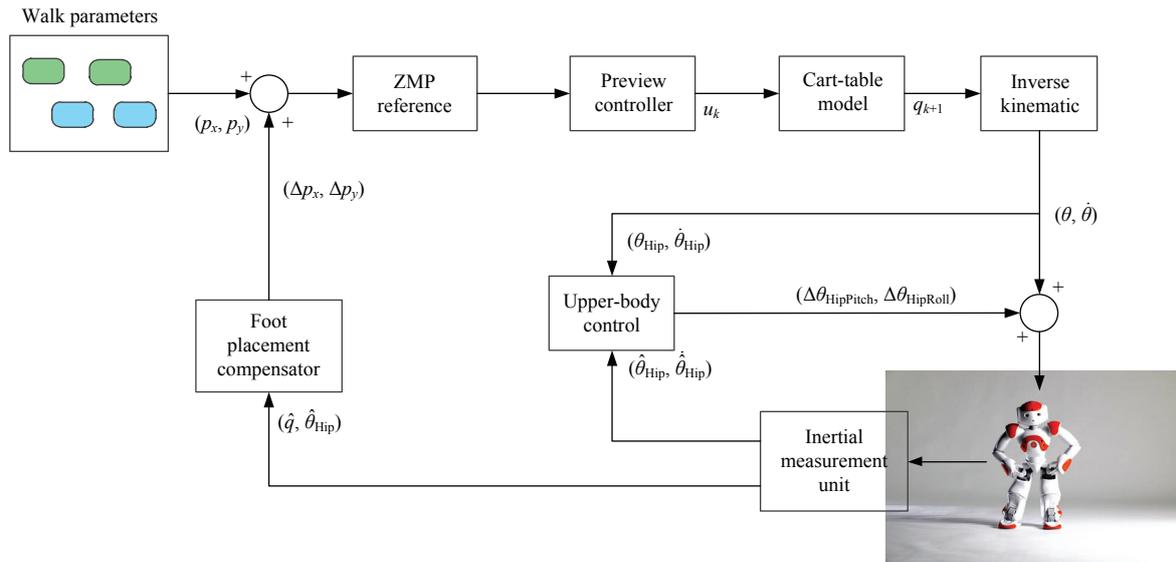


Fig. 5 Diagram of our humanoid walking control with an FPC.

6.1 Restoring balanced walking under unknown perturbations

Firstly, the robot is controlled to walk forward straightly (along the x -axis) on level ground with a stride of 0.1 m when there are no disturbances. In the same situation, the right side of robot is subjected to a thrust, in the opposite direction along the y -axis for almost 10 s. Quite a few steps are needed for the robot to recover. Then a force of 12.26 N is put on the center of the robot for almost 0.7 s. Through the presented PGRL approach, the parameters of FPC can be learned. The termination condition of the iteration process is either when the iteration error is not more than the given error border, or the number of iterations is over the boundary 480. 400 iterations are sustained in this simulation. For each iteration, the best policy performances are presented in Fig. 6. After 400th iteration, the process of learning almost converged. The FPC parameters are listed in Table 1. The matrix of gain coefficient ϕ is random generated and step size factor η set as 0.001. The weight factors of CoM error and the weight factors of torso angle error along the x - and y -axes were set as 0.6 and 0.4, respectively.

In the process of experiment, the inclination attitudes of the upper body can be calculated by merging the data from accelerometer and gyroscope sensors, which are presented as Fig. 7. Because of the external force is applied during the walking process, the body roll reaches

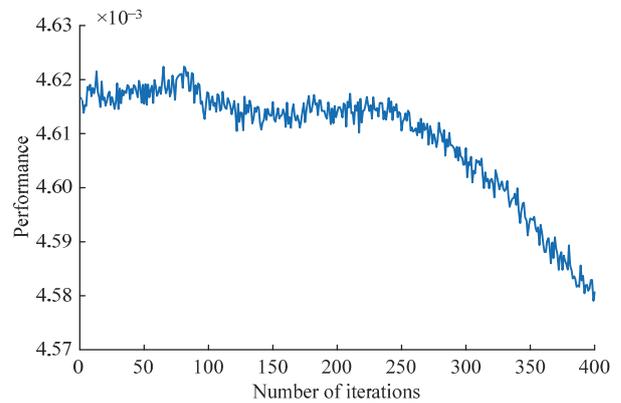


Fig. 6 Iteration performance.

Table 1 Policy gradient method parameters

Parameter	Value
α_x	0.6
α_y	0.4
β_x	0.6
β_y	0.4
ϕ	[0,0,0,0]
η	0.001
ε_j	0.001
N_{iter}	480

the highest point at nearly 12 s and dropped to -0.2 rad while the body pitch curve is almost coincident with the x -axis.

The comparison between a closed-loop with an observer and an open-loop case is presented in Fig. 8.

There is no expressive distinction with the real ZMP curves in these two cases in the situation of no external disturbance. Nonetheless, the preview-control approach with an observer presents smaller fluctuations than that of without an observer. In spite of measurement noises and delay in transmission, the errors in the actual CoM and ZMP curves are opposed to the ideal situations.

Fig. 9 illustrates the reference and actual ZMP trajectories both in the x - and y -directions throughout the walking process. These two plots illustrate that when the impulse is applied, the robot's ZMP rapidly increases and approaches its supporting range. For the purpose of maintaining balance, the robot increases its next stride by referring to the FPC policy and causes the ZMP back to its stable area within next stride. This experiment is presented by snapshots in Fig. 10.

6.2 Walking on uneven terrain

The effectiveness of the walking process on an uneven terrain is proved. In this experiment, the robot walks straightly (x -axis direction) over a square white hard board, whose length is 0.5 m, width is 0.5 m and the height is 0.018 m. This board is a rigid body which means it cannot be deformed. The inclination of the upper body is measured by the sensor merging method in the process of experiment. The comparison between the open- and close-loop (with observer) situations is presented as Fig. 11a. Fig. 11b illustrates the actual and reference ZMP trajectories in the x and y directions

throughout the walking process. Fig. 12 shows snapshots of the experiment.

6.3 Real robot experiment

Two experiments are conducted to verify real robot performance in a humanoid robot with no prior knowledge about the walking terrain. In the first experiment, the robot is stepping and was pushed from the front with a force gauge. The experimenter continually increases thrust to verify the robot's walking performance by applying external disturbance. In the absence of the FPC, the humanoid robot can also withstand some external thrust under the effect of the preview controller and upper body posture controller. The maximum thrust that the robot could withstand is approximately 11 N. After exerting approximately 12.9 N of thrust, the robot falls backward.

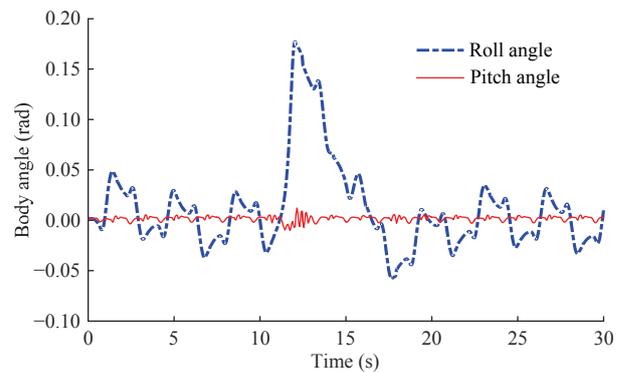


Fig. 7 Upper body attitude.

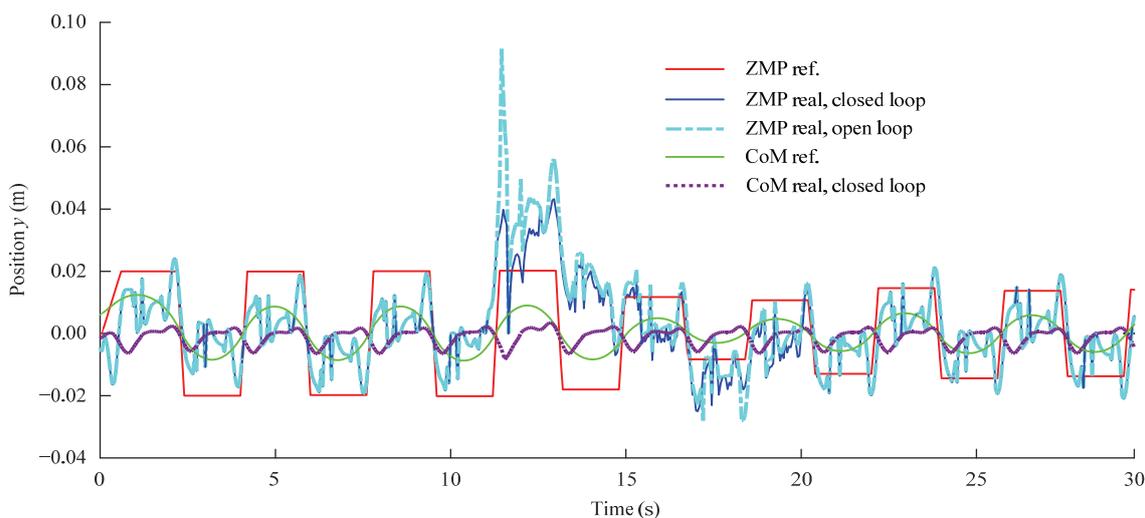


Fig. 8 CoM and ZMP trajectories.

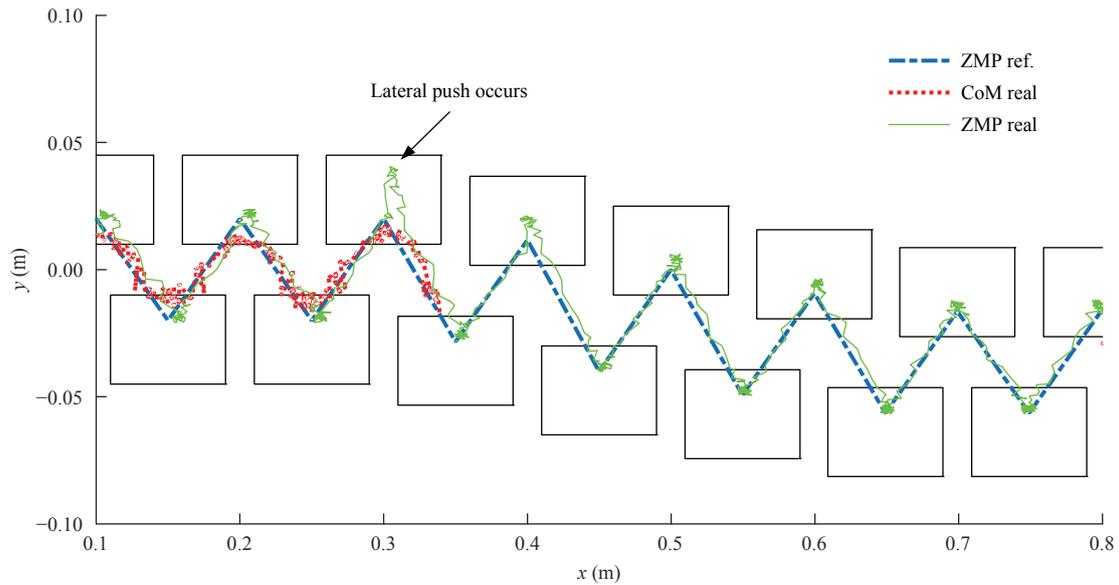


Fig. 9 Real and reference ZMP trajectories in the x - and y -directions.



Fig. 10 Snapshot sequences of the NAO robot recovering from a lateral push.

Fig. 13a shows a snapshot sequence of the experiment without the FPC. With the proposed FPC, when the thrust is more than 12.9 N, the robot can still maintain balance by adjusting its foot placement. The maximum thrust is approximately 15 N. In the method proposed in Ref. [41], the robot can only suffer a momentum impact of 1.13 N·s. The comparative experiments demonstrate that the presented FPC is valid and en-

hances the robustness of humanoid robots considerably. Fig. 13b shows a snapshot sequence of the experiment with the FPC. The experimental result obtained using the proposed FPC is shown in Fig. 14. The external impulse imposes at around 1.1 s, to maintain the walking balance, and the robot increases the next stride size by referring to the FPC policy, allowing the ZMP go back to the normal situation within several steps.

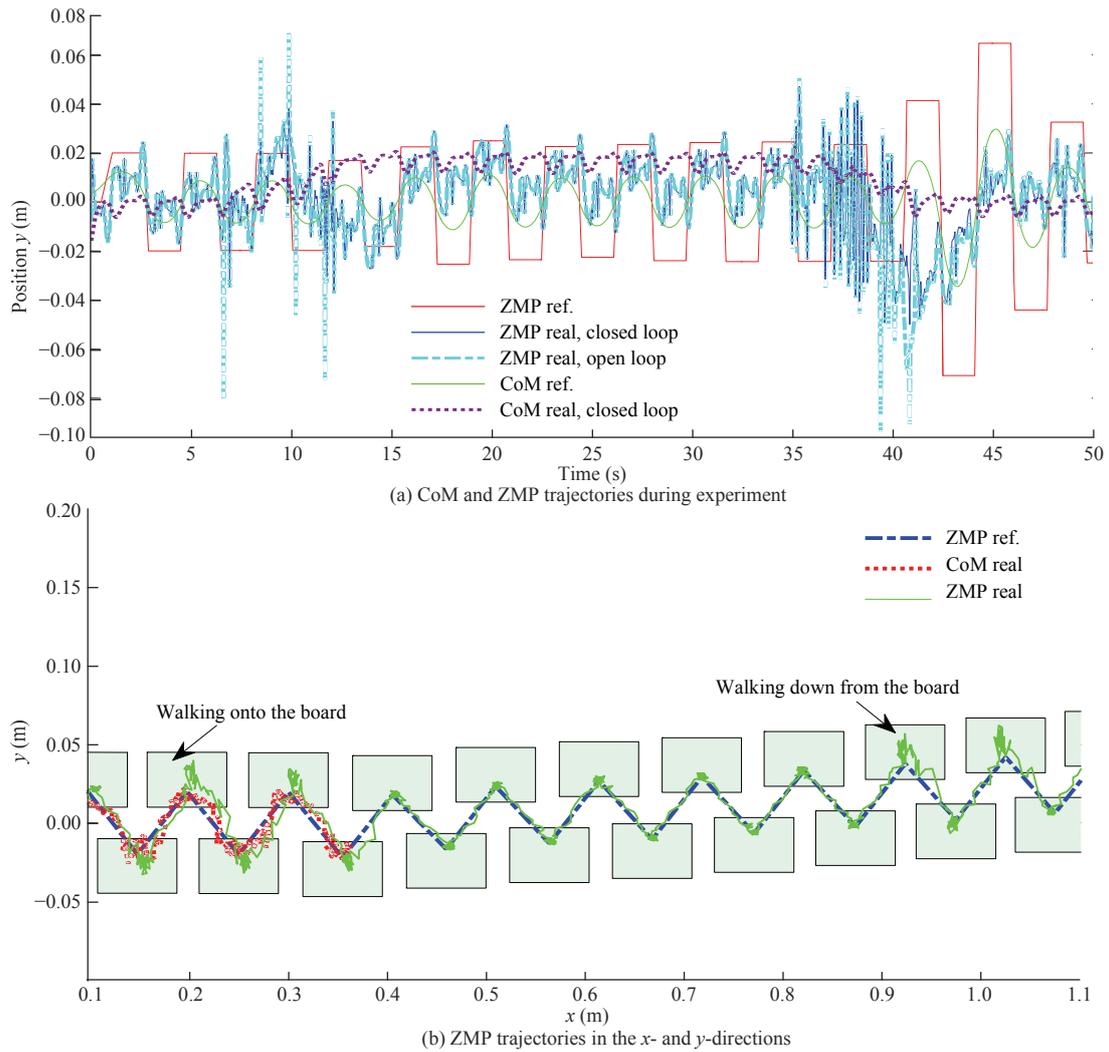


Fig. 11 Real and reference ZMP and CoM trajectories during walking experiment.

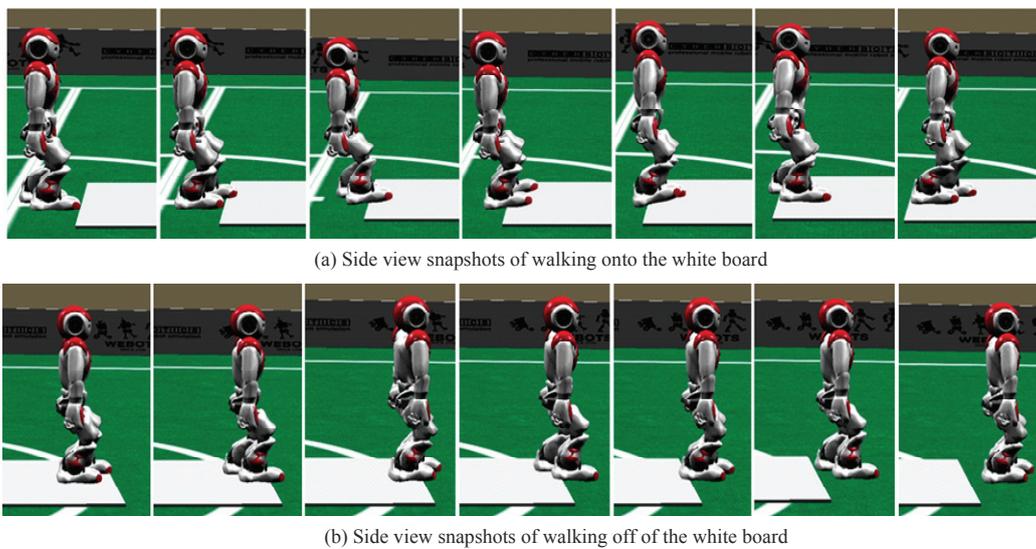


Fig. 12 Snapshot sequences of walking on an uneven terrain.

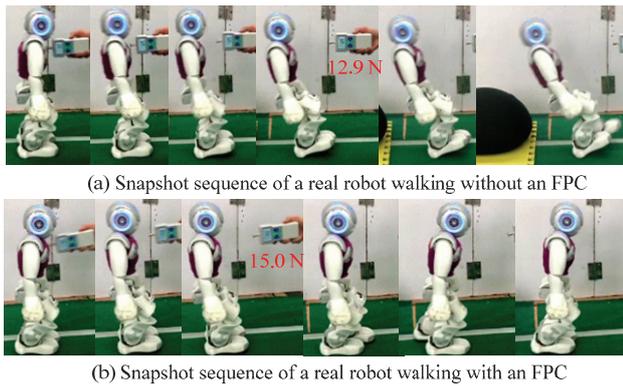


Fig. 13 Snapshot sequence of a real robot recovering from a push.

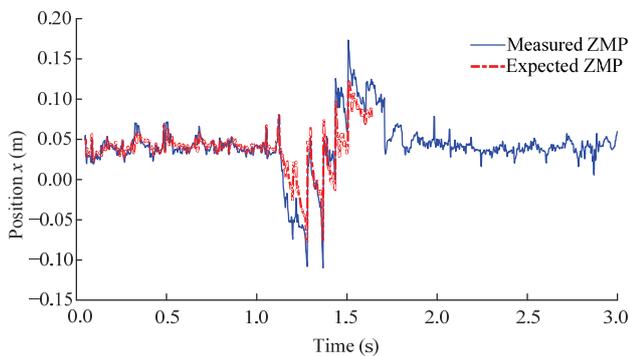


Fig. 14 Real experiment of stepping under unknown perturbations.

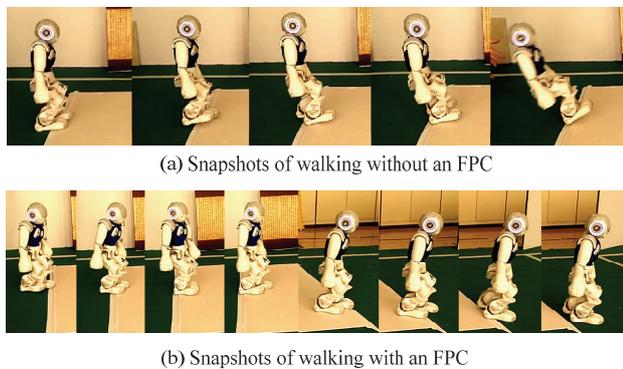


Fig. 15 Snapshot sequences of a real robot experiment showing walking on uneven terrain.

NAO is expected to walk through a hard board in the second experiment. The variation in height from flat terrain to the hard board was approximately 1.0 cm. Fig. 15 presents the snapshots of the experiment. A rebound appears when the NAO lands on its right foot at 3.10 s, and several continuous larger strides are made to reduce the disturbances and then a successful walking process are completed through the platform.

7 Conclusion and future work

In this work, a compensator based on foot positioning for a full-body humanoid was developed to maintain its balance during continuous walking. The foot positioning parameters can be learned according to the presented FPC algorithm, which combined the sensory feedback and the re-planned CoM trajectories and gait pattern by preview control theory. Specifically, in the process of the walking pattern re-planning online, the method PGRL was applied to resist the external disturbances for walking position recovering. Besides, to further enhance the ability of disturbance resistance for humanoid robots, a control strategy based on upper body posture was proposed. The benefits of our approach were verified in different experiments.

The simulated irregular terrain adaptive walking verified the feasibility of the proposed control strategy. Based on the simulation results, the program was dubbed to the real NAO robot to test its adaptive walking performance. Simulations are normally achieved from arbitrary offline designs, rather than from the result of embodied cognitive processes. So, the reality gap problem occurs, meaning that the optimized control parameters obtained from the offline optimization usually do not lead to the same locomotion effect as they do on simulators. One of the main reasons for this phenomenon is that simulation models are only simplifications of the real world. In this work, since ODE cannot provide precise dynamical model and the friction force differences between the simulation and real walking terrains, the gains of the feedback loops were needed to be modulated by trial and error within a very small range.

Future work will focus on the online parameters learning method. A non-parametric regression model with an extended Gaussian process model for online FPC will be added, which can achieve a faster computation speed in real-time, adjustable online model and then reach a better performance of disturbance rejection

Acknowledgment

The authors would like to express thanks to technicians in the Laboratory of Robotics and Intelligent Systems of Tongji University for their assistance during experiments. This work was supported in part by the

National Natural Science Foundation of China (Grant Nos: 61673300 and U1713211), and Basic Research Project of Shanghai Science and Technology Commission (Grant No. 18DZ1200804).

* All supplementary materials are available at <https://doi.org/10.1007/s42235-020-0011-x>.

References

- [1] Zhang J Q, Gao F, Han X L, Chen X B, Han X Y. Trot gait design and CPG method for a quadruped robot. *Journal of Bionic Engineering*, 2014, **11**, 18–25.
- [2] Liu C J, Xia L, Zhang C Z, Chen Q J. Multi-layered CPG for adaptive walking of quadruped robots. *Journal of Bionic Engineering*, 2018, **15**, 341–355.
- [3] Bai K Q, Luo M Z, Li T, Wu J. The impulse excitation joint servo drive design and adaptive backstepping control of humanoid robots. *Journal of Bionic Engineering*, 2018, **15**, 114–125.
- [4] Huang Q, Yokoi K, Kajita S, Kaneko K, Arai H, Koyachi N, Tanie K. Planning walking patterns for a biped robot. *IEEE Transactions on Robotics and Automation*, 2001, **17**, 280–289.
- [5] Seven U, Akbas T, Fidan K C, Erbatur K. Bipedal robot walking control on inclined planes by fuzzy reference trajectory modification. *Soft Computing*, 2012, **16**, 1959–1976.
- [6] Kajita S, Kanehiro F, Kaneko K, Yokoi K, Hirukawa H. The 3D Linear Inverted Pendulum Mode: A simple modeling for a biped walking pattern generation. *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Maui, HI, USA, 2001, 239–246.
- [7] Kajita S, Morisawa M, Miura K, Nakaoka S I, Harada K, Kaneko K, Yokoi K. Biped walking stabilization based on linear inverted pendulum tracking. *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Taipei, Taiwan, 2010, 4489–4496.
- [8] Graf C, Röfer T. A center of mass observing 3D-LIPM gait for the RoboCup standard platform league humanoid. *Robot Soccer World Cup*, 2011, 102–113.
- [9] Graf C, Rofer T T. A closed-loop 3D-LIPM gait for the RoboCup standard platform league humanoid. *Proceedings of 5th Workshop Humanoid Soccer Robot & Humanoids*, Nashville, USA, 2010, 18–22.
- [10] Kajita S, Kanehiro F, Kaneko K, Fujiwara K, Harada K, Yokoi K, Hirukawa H. Biped walking pattern generation by using preview control of zero-moment point. *Proceedings of IEEE International Conference on Robotics and Automation (ICRA)*, Taipei, Taiwan, 2003, 1620–1626.
- [11] Motoi N, Suzuki T, Ohnishi K. A bipedal locomotion planning based on virtual linear inverted pendulum mode. *IEEE Transactions on Industrial Electronics*, 2008, **56**, 54–61.
- [12] Yu Z G, Chen X C, Huang Q, Zhang W, Meng L B, Zhang W M, Gao J Y. Gait planning of omnidirectional walk on inclined ground for biped robots. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2015, **46**, 888–897.
- [13] Kim M, Collins S H. Once-per-step control of ankle push-off work improves balance in a three-dimensional simulation of bipedal walking. *IEEE Transactions on Robotics*, 2017, **33**, 406–418.
- [14] Yu Z G, Zhou Q Q, Chen X C, Li Q Q, Meng L B, Zhang W M, Huang Q. Disturbance rejection for biped walking using zero-moment point variation based on body acceleration. *IEEE Transactions on Industrial Informatics*, 2018, **15**, 2265–2276.
- [15] Castano J A, Li Z, Zhou C, Tsagarakis N, Caldwell D. Dynamic and reactive walking for humanoid robots based on foot placement control. *International Journal of Humanoid Robotics*, 2016, **13**, 1550041.
- [16] Hu K, Ott C, Lee D. Learning and generalization of compensative zero-moment point trajectory for biped walking. *IEEE Transactions on Robotics*, 2016, **32**, 717–725.
- [17] Mason S, Rotella N, Schaal S, Righetti L. An MPC walking framework with external contact forces. *Proceedings of International Conference on Robotics and Automation (ICRA)*, Brisbane, Australia, 2018, 1785–1790.
- [18] Chen X C, Yu Z G, Zhang W M, Zheng Y, Huang Q, Ming A G. Bioinspired control of walking with toe-off, heel-strike, and disturbance rejection for a biped robot. *IEEE Transactions on Industrial Electronics*, 2017, **64**, 7962–7971.
- [19] Liu C J, Wang D W, Chen Q J. Central pattern generator inspired control for adaptive walking of biped robots. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2013, **43**, 1206–1215.
- [20] Liu C J, Wang D W, Goodman E D, Chen Q J. Adaptive walking control of biped robots using online trajectory generation method based on neural oscillators. *Journal of Bionic Engineering*, 2016, **13**, 572–584.
- [21] Saputra A A, Botzheim J, Sulistijono I A, Kubota N. Biologically inspired control system for 3-D locomotion of a humanoid biped robot. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2015, **46**, 898–911.
- [22] Wang Y, Xue X H, Chen B F. Matsuoka's CPG with desired rhythmic signals for adaptive walking of humanoid robots. *IEEE Transactions on Cybernetics*, 2018, **50**, 613–626.

- [23] Wieber P B. Trajectory free linear model predictive control for stable walking in the presence of strong perturbations. *Proceedings of 6th IEEE-RAS International Conference on Humanoid Robots*, Genova, Italy, 2006, 137–142.
- [24] Diedam H, Dimitrov D, Wieber P B, Mombaur K, Diehl M. Online walking gait generation with adaptive foot positioning through linear model predictive control. *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Nice, France, 2008, 1121–1126.
- [25] Herdt A, Diedam H, Wieber P B, Dimitrov D, Mombaur K, Diehl M. Online walking motion generation with automatic footstep placement. *Advanced Robotics*, 2010, **24**, 719–737.
- [26] Nishiwaki K, Kagami S. Strategies for adjusting the ZMP reference trajectory for maintaining balance in humanoid walking. *Proceedings of IEEE International Conference on Robotics and Automation*, Alaska, USA, 2010, 4230–4236.
- [27] Feng S, Xinjilefu X, Atkeson C G, Kim J. Robust dynamic walking using online foot step optimization. *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Daejeon, Korea, 2016, 5373–5378.
- [28] Kryczka P, Kormushev P, Tsagarakis N G, Caldwell D G. Online regeneration of bipedal walking gait pattern optimizing footstep placement and timing. *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Hamburg, Germany, 2015, 3352–3357.
- [29] Missura M, Behnke S. Online learning of foot placement for balanced bipedal walking. *Proceedings of IEEE-RAS International Conference on Humanoid Robot*, Madrid, Spain, 2014, 322–328.
- [30] Missura M, Behnke S. Gradient-driven online learning of bipedal push recovery. *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Hamburg, Germany, 2015, 387–392.
- [31] Xu T, Chen Q J, Cai Z Q. Rebalance strategies for humanoids walking by foot positioning compensator based on adaptive heteroscedastic SpGPs. *Proceedings of International Conference on Robotics and Automation (ICRA)*, Shanghai, China, 2011, 563–568.
- [32] Liu C J, Xu T, Wang D W, Chen Q J. Active balance of humanoids with foot positioning compensation and non-parametric adaptation. *Robotics and Autonomous Systems*, 2016, **75**, 297–309.
- [33] Kohl N, Stone P. Policy gradient reinforcement learning for fast quadrupedal locomotion. *Proceedings of International Conference on Robotics and Automation (ICRA)*, LA, USA, 2004, **3**, 2619–2624.
- [34] Cherubini A, Giannone F, Iocchi L, Lombardo M, Oriolo G. Policy gradient learning for a humanoid soccer robot. *Robotics and Autonomous Systems*, 2009, **57**, 808–818.
- [35] Kajita S, Hirukawa H. *Humanoid Robot*, Ohmsha Ltd, Tokyo, Japan, 2005.
- [36] Katayama T, Ohki T, Inoue T, Kato T. Design of an optimal controller for a discrete-time system subject to previewable demand. *International Journal of Control*, 1985, **41**, 677–699.
- [37] Liao F C, Chen P. Optimal preview control based on state observers for linear discrete-time systems. *Journal of University of Science and Technology*, 2014, **36**, 390–398. (in Chinese)
- [38] Jadidi M G, Hashemi E. Optimal preview control of the nao biped robot using a ukf-based state observer. *Proceedings of IEEE International Conference on Advanced Intelligent Mechatronics (AIM)*. Banff, Alberta, Canada, 2016, 52–57.
- [39] Shimmyo S, Sato T, Ohnishi K. Biped walking pattern generation by using preview control based on three-mass model. *IEEE Transactions on Industrial Electronics*, 2012, **60**, 5137–5147.
- [40] Zhou C X, Li Z B, Wang X, Tsagarakis N, Caldwell D. Stabilization of bipedal walking based on compliance control. *Autonomous Robots*, 2016, **40**, 1041–1057.
- [41] Urbann O, Hofmann M. A reactive stepping algorithm based on preview controller with observer for biped robots. *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Daejeon, Korea, 2016, 5324–5331.