Available online at www.sciencedirect.com
ScienceDirect
Engineering Applications of Artificial Intelligence 21 (2008) 1027–1034
Engineering Applications of ARTIFICIAL INTELLIGENCE
www.elsevier.com/locate/engappai

Passive dynamic walker controller design employing an RLS-based natural actor–critic learning algorithm

Baeksuk Chu[a], Daehie Hong[a,*], Jooyoung Park[b], Jae-Hun Chung[c]

[a]Department of Mechanical Engineering, Korea University, 5-1, Anam-dong, Sungbuk-gu, Seoul 136-701, Republic of Korea
[b]Department of Control and Instrumentation Engineering, Korea University, Jochiwon, Chungnam 339-700, Republic of Korea
[c]Department of Mechanical Engineering, Stevens Institute of Technology, Castle Point on Hudson, Hoboken, NJ 07030, USA

Received 21 March 2007; received in revised form 5 November 2007; accepted 11 November 2007
Available online 26 December 2007

Abstract

A passive dynamic walker belongs to a class of bipedal walking robots that are able to walk stably down a small decline without using any actuators. The purpose of this research is to design a controller in order to build actuated robots capable of walking on a flat terrain based on passive dynamic walking. To achieve this objective, a control algorithm was used based on reinforcement learning (RL). The RL method is a goal-directed learning of a mapping from situations to actions without relying on exemplary supervision or complete models of the environment. The goal of the RL method is to maximize a reward, which is an evaluative feedback from the environment. In the process of constructing the reward of the actuated passive dynamic walker, the control objective, which is stable walking on level ground, is directly included. In this study, an RL algorithm based on the actor–critic architecture and the natural gradient method is applied. Also, the recursive least-squares (RLS) method was employed for the learning process in order to improve the efficiency of the method. The control algorithm was verified with computer simulations based on the eigenvalue analysis for stable locomotion.

Keywords: Passive dynamic walker; Reinforcement learning (RL); Actor–critic architecture; Natural gradient; Recursive least-squares (RLS)

1. Introduction

Mainstream bipedal robots that have achieved impressive accomplishments such as Honda Motor Company's humanoid robot, ASIMO, are usually based on statically stable gait and a precise joint angle control scheme (Hirai et al., 1998). The mainstream robots actively control every joint angle at all times with substantial real-time computation and need actuators with high frequency and power. Hence, their control paradigms are considerably different from human walking and are not desired in the aspect of rehabilitating human locomotion (Collins et al., 2005). In the late 1980s, a class of walking robots that were able to walk stably down a small decline without any actuation or control was introduced by McGeer (1990). These machines are called 'passive dynamic walker', and, through elaborate kinematical design, can generate stable and very energy efficient walking. Moreover, employing knee-bending motions and balancing through arm swing, recently developed passive dynamic walkers are able to walk with an amazingly human-like gait (Collins et al., 2001, 2005). The mechanical design of the robot treated in this study is based on the concept of such passive dynamic walking. The objective of this research is to design a controller in order to build actuated robots capable of walking stably on a flat terrain on the framework of passive dynamic walking.

Some problems have been reported about applying analytic control designs to a dynamical bipedal system. First, it is an under-actuated system. Namely, because the number of actuators is less than the degrees of freedom to be controlled, the control is very complicated. Second, uncertainties in contact with the ground and nonlinear friction in the joints make the model-based design difficult. Lastly, it has dynamic discontinuities caused by collisions with the ground. Due to these characteristics, the control of a passive dynamic walker has been known as very challenging task to accomplish. To overcome such problems, an

*Corresponding author. Tel.: +82 2 3290 3369; fax: +82 2 926 9290.
E-mail address: dhhong@korea.ac.kr (D. Hong).

0952-1976/$ - see front matter
doi:10.1016/j.engappai.2007.11.001

reinforcement learning (RL)-based learning method is employed as the control algorithm used in this study, which is a goal-directed learning of a mapping from situations to actions without relying on exemplary supervision or complete models of the environment. In RL methodologies, a general model such as a neural network can be selected as a controller that has computational learning capability through direct interaction with environment. The goal of the RL method is to maximize a reward or reinforcement signal that is an evaluative feedback from the environment. The control objective of stable walking on level ground is incorporated directly into the process of constructing the reward for a passive dynamic walker.

RL has been an active research area in machine learning, control engineering, and so on (Kaelbling et al., 1996; Sutton and Barto, 1998). Among the many categories of RL, this study is based on the actor–critic learning algorithm which is sometimes called adaptive–heuristic–critic (AHC) learning architecture (Kaelbling et al., 1996). In this class of learning structure, the controller is divided into two components: the critic (evaluation) module and the actor (control) module. Both modules have their own learning processes, and make the controller approach the optimal performance with respect to the reward. In this paper, the actor adjustment is determined by the 'natural gradient' method whose learning performance is reportedly better than other gradient optimization algorithms. And the critic module is adjusted by a 'recursive least-squares (RLS)'-based estimation algorithm in order to improve the efficiency of the use of data. Such a combination in actor–critic architecture is called an 'RLS-based natural actor–critic algorithm (Park et al., 2005)'. This algorithm, for the first time, is applied to control the actuated version of the passive dynamic walker in this paper. An RLS-based natural actor–critic algorithm is modified to fit to the target system and corresponding programming and simulations are performed. With a reasonable evaluation method, a superior result compared to the previous control research is obtained.

This paper is organized as follows. In Section 2, the target passive dynamic walker is briefly introduced. In Section 3, the basic concepts of the RL and the actor–critic learning algorithms based on the natural gradient and RLS estimation scheme are demonstrated for control design for the walking robot. Then, in Section 4, the simulation results are discussed. Finally, Section 5 contains concluding remarks.

2. Passive dynamic walker

The passive dynamic walker as shown in Fig. 1 represents the simplest configuration of its kinematics in which stable dynamic walking in three dimensions can be maintained (Tedrake, 2004). It has only a single passive pin joint at the hip. When it is located at a high place with a small ramp and a small push is given sideways, stable walking down begins only with the help of gravity. The walker will rock onto a single stance leg, allowing the opposite leg to leave the ground and swing forward down the ramp. Upon landing, the robot rocks onto the opposite foot, and the cycle iterates.

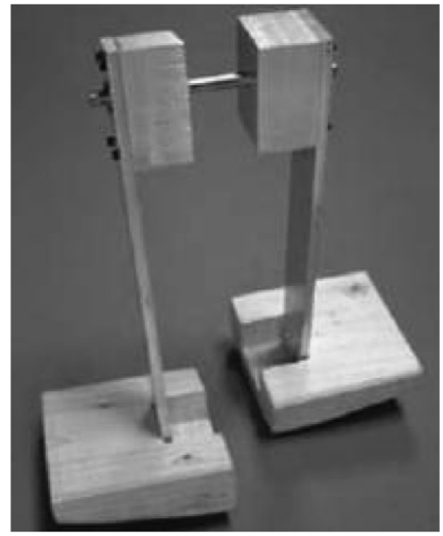

Fig. 1. The 3D passive dynamic walker (Tedrake, 2004).

The locomotion of the passive dynamic walker is composed of sagittal plane dynamics and frontal plane dynamics. A sagittal plane model is just about the robot's walking down the ramp from the side view, and a frontal plane rocking model is needed for generating a foot clearance from the ground when the swing leg passes forward the stance leg. The dynamics of the sagittal plane has a stable property and can be controlled independently using a simple velocity control algorithm (Tedrake, 2004). Also, noting that the ground contact of legs takes place with half period of the frontal plane oscillations, the sagittal plane model is easily coupled to the frontal plane model. Therefore, to solve the problem regarding this robot's locomotion, we first focus our attention on stabilizing the oscillation in the frontal plane.

For the purpose of modeling the dynamics in the frontal plane as shown in Fig. 2, it is assumed that the robot always touches the ground at only one point and there is no slip during rolling of the feet. With respect to the body angle φ, the equations of motion of the planar system are represented as

$$H(\varphi)\ddot{\varphi} + C(\varphi, \dot{\varphi})\dot{\varphi} + G(\varphi) = 0. \tag{1}$$

When $|\varphi| > \omega$, ground contact occurs in the curved area of the foot, and the dynamics can be written as

$$H(\varphi) = I + ma^2 + m\rho^2 - 2m\rho a \cos\varphi, \tag{2}$$

$$C(\varphi, \dot{\varphi}) = m\rho a\dot{\varphi} \sin\varphi, \tag{3}$$

$$G(\varphi) = mga \sin\varphi, \tag{4}$$