Control Engineering Practice 19 (2011) 89–99

Contents lists available at ScienceDirect

Control Engineering Practice

journal homepage: www.elsevier.com/locate/conengprac

Real time implementation of CTRNN and BPTT algorithm to learn on-line biped robot balance: Experiments on the standing posture

Patrick Hénaff [a,*], Vincent Scesa [b], Fethi Ben Ouezdou [b], Olivier Bruneau [b]

[a] ETIS, UMR 8051, CNRS-ENSEA-UCP, University of Cergy-Pontoise, Site de ST Martin, 2 rue A. Chauvin, 95302 Cergy-Pontoise Cedex, F-9500, France
[b] LISV, University of Versailles Saint-Quentin, France

ARTICLE INFO

Article history:
Received 7 December 2009
Accepted 6 October 2010
Available online 30 October 2010

Keywords:
Neural control
Learning algorithms
CTRNN
Real-time systems
Robotics
Biped robot

ABSTRACT

This paper describes experimental results regarding the real time implementation of continuous time recurrent neural networks (CTRNN) and the dynamic back-propagation through time (BPTT) algorithm for the on-line learning control laws. Experiments are carried out to control the balance of a biped robot prototype in its standing posture. The neural controller is trained to compensate for external perturbations by controlling the torso's joint motions. Algorithms are embedded in the real time electronic unit of the robot. On-line learning implementations are presented in detail. The results on learning behavior and control performance demonstrate the strength and the efficiency of the proposed approach.

1. Introduction

Technological developments have enabled us to build robots with morphologies that are inspired by animals or humans. Therefore, the most recent humanoid robots are technologically complex systems, with an extremely high level of mechanical and electronic integration. They are equipped with complete perceptive systems that enable them to interact with human beings and to move in an environment built for human life. One of the most important difficulties in controlling humanoid robots is maintaining balance during walking or standing. One solution to this problem is to design controllers based on the zero moment point (ZMP) theory (Vukobratovic, 2004). Another method is to design controllers using bio-inspired approaches, i.e., with some capabilities of adaptation and training, leading to the acquisition of reflexes.

Using biologically inspired architectures such as neural networks that are able to learn the "correct" control of the robot's equilibrium is a promising approach. For this purpose, several neural controller-based approaches have been proposed in the past. Cerebellar model articulation controllers (CMAC) proposed in 1975 by Albus (1975) are still studied in the control of legged robots. Recent studies deal with their modeling and generalization properties (Horvath & Szabo, 2007) or with their connections to other approaches like fuzzy logic (Su, Lee, & Wang, 2006) or computed torque control (Lin & Chen, 2007). CMAC have been used to control the balance of biped robots (Kun & Miller, 1996) or for robust dynamic walking in simulation (Lin & Chen, 2007) and for biped robot experiments (Sabourin & Bruneau, 2005).

Recurrent neural networks (i.e., dynamic neural nets) have been extensively studied in the control of complex systems for many years (Marcua, Köppen-Seligerb, & Stücher, 2008; Song & Tahk, 2001). These artificial neural networks have also been used to design stable walking gaits for biped robots (Wu, Song, & Yang, 2007). Several approaches were based on evolutionary synthesis (Fukuda, Komata, & Arakawa, 1997), neural oscillators (Taga, Yamaguchi, & Shimizu, 1991; Geng, Porr, & Wörgötter, 2006), and central pattern generators (Nakanishi et al., 2004; Righetti & Ijspeert, 2006). Recently, researchers have used RNN as predictive compensator (Mizunoa, Kurodaa, Okazakib, & Ohtsu, 2007) or tracking controllers with self-constructing properties. Self-constructing algorithms are very interesting approaches because they allow for optimizing on-line the neuronal controller architecture in order to insure the best control performance. In Gao and Er (2003), a self-constructing fuzzy neural controller was proposed for the tracking control of a simulated planar robot manipulator with two degrees of freedom. In Hsu (2009) a simple growing and pruning algorithm applied to a recurrent neural network has been tested in experimentation to control one degree of freedom of a moving table with a linear ceramic motor system.

Many studies on dynamic neural controllers of robots have focused on continuous-time recurrent neural networks (CTRNN) due to their ability to be universal approximators (Beer, 2006). CTRNNs have been used for bio-inspired control because of their abilities to reproduce the full qualitative range of nerve cell

* Corresponding author. Tel.: +33 1 34 25 28 51; fax :+33 1 34 25 28 39.
E-mail addresses: patrick.henaff@u-cergy.fr (P. Hénaff), fethi.benouezdou@uvsq.fr (F. Ben Ouezdou), bruneau@lisv.uvsq.fr (O. Bruneau).

0967-0661/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.conengprac.2010.10.002

phenomenology (Beer, 2006). They make it possible to show adaptivity properties based on homeostatic plastic mechanisms (Hoinville & Hénaff, 2004; Williams, 2007). Moreover, they may be a fine model for generating adaptive behavior because they can learn with the back-propagation through time algorithm (BPTT) (Pineda, 1987; Pearlmutter, 1995; Rumelhart, Hinton, & Williams, 1986; Robinson & Fallside,1987; Werbos, 1990). Unfortunately, these algorithms are complex to implement in real time application, especially for the on-line training of real robots because of the shrinking gradient problem of recurrent neural networks (this gradient instability problem was first studied in Scesa, Hénaff, Ouezdou, & Namoun, 2006).

There are still only a few recent in-depth studies on real time implementation, especially with the concept of learning the equilibrium reflex for a biped robot on-line. The scientific objective of the work presented in this paper is to perform an in-depth analysis of the real time performance of continuous time recurrent neural network controllers for learning balance reflexes for a biped robot. In particular, robot abilities to learn on-line with real time constraints are investigated.

Several contributions in this area have already been published (see Scesa, Mohamed, Henaff, & Ouezdou, 2005). This paper focuses on describing the real time implementation of the learning algorithms embedded into the robot control unit. Some promising new experimental results are also reported.

This paper is organized as follows. The second section deals with the learning models based on the CTRNN and BPTT algorithms. The third section presents the fundamental principles of CTRNN and BPTT. In the fourth section, the experimental biped platform called ROBIAN is described. The real time implementation of the learning algorithm based on CTRNNs and the back-propagation through time algorithm is detailed in Section 5. The sixth section, describes the experimental results in two subparts: a feasibility test for the embedded learning algorithm on controlling the ROBIAN biped torso and a new approach of on-line learning of the equilibrium reflex. Finally, the last section presents the conclusions and potential further developments stemming from this work.

2. Learning equilibrium with CTRNN

It is well known that the human torso attempts to stabilize the whole body when walking (Hyon, 2009; Kubica, Wang, & Winter, 2001; Setiawan, Hyon, Yamaguchi, & Takanishi, 1999). It is also known that biped robot balance can be considered to be a global behavior because the upper part (the torso) and the lower part (the legs) of the robot interact with and disturb each other (see Fig. 1) during walking. This internal interaction can be taken into account in the synthesis of the biped robot controller. For example, Morimoto et al. (2006) showed that a biped robot can stop and walk using simple sinusoidal desired trajectories with their phases adjusted by a coupled oscillator model.

During the walking and halting phases, the robot's balance can be affected by perturbations like external forces applied to its body, including the upper part. One way to control the balance of the robot is to drive the movements of the upper part (i.e., the torso) in order to minimize the perturbations (forces and moments) exerted on the lower part of the robot.

To measure this perturbation, the robot is equipped with a six-force sensor fixed between the torso and the lower part (see Fig. 1). Then, if the forces and moments measured by the sensor are close to zero, the equilibrium of the robot is assumed to be controlled (Mohammed, Gravez, & Ouezdou, 2004).

To achieve this goal, the purpose of this experiment is to learn on-line, using a CTRNN, how to control the robot torso in order to reduce the external perturbations measured by the sensor. This learning is carried out with the biped robot in the standing posture when unknown external forces are applied to its torso.

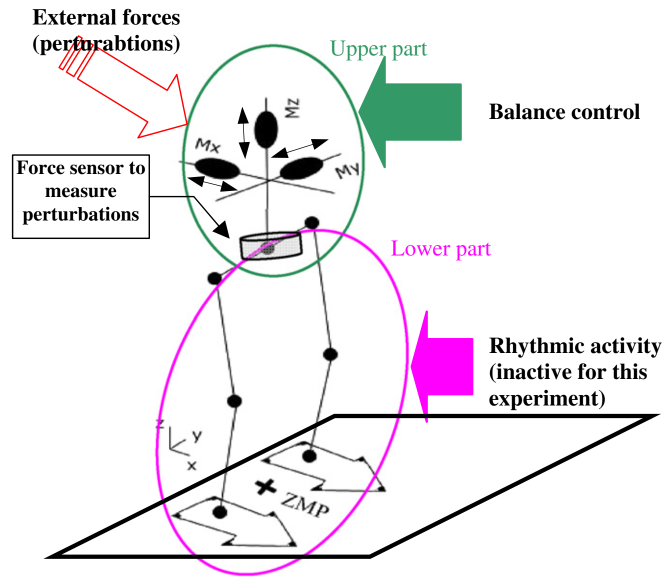

Fig. 1. Control of the biped robot's balance: perturbations caused by unknown external forces can be measured with a force sensor placed between upper and lower part.

The control of robot balance must take into account the real features of the mechanism and the external phenomena that are not modeled in the simulation (friction, motor properties, attrition, ground slope and passive prosthesis use…). Information on these phenomena is often only available through their effects on the total energy of the system. Hence, taking into account these phenomena into the equilibrium control is a difficult task. To avoid this obstacle, the controller should be able to adapt its behavior in real time, following a cost function that incorporates information on these phenomena. Moreover, in an optimal control approach, time variations must be included in the adaptation process.

To meet these conditions, dynamic recurrent neural networks with back-propagation through a time learning process was chosen as the most appropriate network. The proposed learning control architecture is shown in Fig. 2. The outputs of the network are positions (X,Y,Z) that the torso mechanism then has to reach. Thus, at each time step, the net modifies the actuator speed and accelerations are generated. Consequently, forces between the upper and lower parts of the robot will be produced to compensate the external perturbations.

The inputs should show the network the current state of the system. They must be enough representatives to compute a correct control response. They consist of the measured components provided by the 6-forces sensor and the positions of the current motors given by incremental encoders.

To carry out the on-line learning of an optimal control, the parameters of the net must be adapted in real time while the system is running. Back-propagation through time (BPTT), with its ability to integrate the error in the network at each instant is an appropriate solution to solve the parameter adaptation problem. In BPTT, the network is first unrolled in time creating a layer per time step. Then, this algorithm back-propagates the output error on these virtual temporal layers as the classical back-propagation algorithm does on existing ones. The result is the computation of the error gradient in the network and its integration in time. Then, the parameters are modified by the gradient descent algorithm, and the response of the net approaches the