Engineering Applications of Artificial Intelligence 55 (2016) 103–118

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

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

Three-level hierarchical model-free learning approach to trajectory tracking control

Mircea-Bogdan Radac [a], Radu-Emil Precup [a,b,*]

[a] Department of Automation and Applied Informatics, Politehnica University of Timisoara, Bd. V. Parvan 2, 300223 Timisoara, Romania
[b] School of Engineering, Edith Cowan University, 270 Joondalup Dr., Joondalup, WA 6027, Australia

ARTICLE INFO

Article history:
Received 27 September 2015
Received in revised form
16 May 2016
Accepted 18 June 2016

Keywords:
Aerodynamic system
Model-free Iterative Learning Control
Neural networks
Reference trajectory tracking
Virtual Reference Feedback Tuning

ABSTRACT

This paper suggests a novel three-level model-free hierarchical learning approach that solves the reference trajectory tracking problem for control systems (CSs). The new approach consists of the low level, the intermediate level and the high level, it relies on past memorized optimal input output execution patterns and adaptively merges them using a similarity measure. The low level feedback control is carried out in a novel model-free framework using a neural network (NN) controwller tuned by Virtual Reference Feedback Tuning (VRFT) in order to linearize the closed-loop CS and to match a linear reference model. The NN controller is tuned in two phases, an offline one and an online one. Nonlinear Model Predictive Control (NMPC) is first employed in the novel offline tuning phase. The online tuning phase makes next use only of the process sign in the dynamic back-propagation mechanism that updates the NN parameters. After the NN controller is trained and the feedback CS is fixed, the optimal execution patterns (input/output patterns) are defined in terms of optimal control problems, which balance control accuracy and control effort. The input/output patterns are formulated over a feedback CS and are solved in a model-free Iterative Learning Control (ILC) framework for linear time-invariant systems at the intermediate level. Once the optimal executions are learned at the intermediate level over the feedback CS, they are stored in a database (DB). Then each time a new trajectory is to be tracked as required by the high level planner, similar patterns of executions are looked-up in the DB and are merged by the weighted average sum of most similar patterns resulted from a sort algorithm using a distance metric. The proposed approach is tested on the position control of a Single Input-Single Output nonlinear aerodynamic control system and shows trajectory tracking performance improvement with respect to the case when no learnt experience is used.

1. Introduction

As the control systems (CS) become more and more complex and they need to cope with new situations to which they were not exposed before, the cognitive control capabilities are more and more demanding. The biological systems seem to approach the problem of solving complex tasks not explicitly mathematically but by combining the accumulated knowledge stored in memory and formulated as primitives, referred to also as strategies (Mussa-Ivaldi and Solla, 2004). Therefore, by combining primitives the living organisms are capable to carry out more complex maneuvers. The living organisms next subsequently add this capability to their current knowledge base, and an extension of the knowledge base is achieved. In this sense, several abilities are needed for the brain of biological systems, such as:

– Strategy projection to achieve the goal, this ability is associated with reasoning and planning.
– Ability to decompose the strategy into well-known strategies, this also represents a planning ability.
– Necessity of storing the learned strategies, this ability is associated with the memory.
– Feedback to improve new strategies by repeated trials, the feedback abilities are associated with learning.

The brain acts as a high level hierarchical control planner and supervisor that coordinates the low level control at the neuromuscular system. Learning therefore occurs at every hierarchical control level under different forms. This learning is mainly done by fusing the information from visual, tactile, auditory and olfactory sensors.

The classes of control structures and algorithms in motor control learning of the biological systems are: predictive or feedforward control, reactive control, which uses the sensor information to update control, and biomechanical control. All these

* Corresponding author at: Department of Automation and Applied Informatics, Politehnica University of Timisoara, Bd. V. Parvan 2, 300223 Timisoara, Romania.
E-mail addresses: mircea.radac@upt.ro (M.-B. Radac), radu.precup@upt.ro (R.-E. Precup).

http://dx.doi.org/10.1016/j.engappai.2016.06.009
0952-1976/

structures and algorithms are addressable within optimal feedback control frameworks (Wolpert et al., 2011). Several types of motor learning processes that are acknowledged in practice are: error-based learning, reinforcement learning and use-dependent learning (Wolpert et al., 2011). The error-based learning exploits the gradients of the error with respect to changes in the motor control signals, while the use-dependent learning deals with continuous changing of the motor system due to pure repetitions; hence, the variability of the specific executed task is reduced, but also a bias is induced towards this trained direction such that when a slightly different motion is performed, relearning is needed.

The aforementioned types of learning occur in a diversity of situations when, for example, the specifications change in terms of tasks/motions have to be executed at different speeds, with different accuracies or in changing environments. The learning can be analyzed in comparison with the combination of learning and adaptation that occurs in classical two-degree-of-freedom control systems structures, both in the feedback structures (in closed-loop) or in the feed-forward control ones, either in iterative control approaches or in adaptive control ones. The major characteristic of these types of learning behavior is however, that the mathematical model (or the internal representation) of the environment is not explicitly used to back-propagate the error information to the corrective actions. Although the learning and adaptation are formalized in theory, the actual achievement in the neuromuscular system is still unknown (Wolpert et al., 2011).

Another learning situation that occurs in biological systems – associated with high level learning and planning capability – is the one in which, in the same environment, a new task is to be performed and was never seen before as, for example, a new different hand reaching motion. The brain seems to be able to predict execution plans for the limb that delivers nearly optimal executions in terms of motion performance. In doing so, it seems to be able to mix and merge different already optimized/learned motion patterns (strategies) that are stored in the memory. This task is again a black-box one by not being explicitly computational on one hand and by also being done outside conscious awareness on the other hand. This concept corresponds to one formal representation in motor learning called primitive-based learning (Wolpert et al., 2011).

The combinations of the above mentioned learning concepts are the motivation and the main contribution for this paper's work, which proposes a three-level model-free hierarchical learning approach that solves the reference trajectory tracking problem for control systems (CSs). This new approach, which consists of the low level, the intermediate level and the high level, is focused on the replication of some of the learning, planning and prediction mechanisms presented above.

The low level feedback controller tuning is carried out using a novel neural network (NN)-Virtual Reference Feedback Tuning (NN-VRFT) approach for which the tuning is carried out in two steps in a Direct Model Reference Adaptive Control (DMRAC) setting. The first tuning phase of the NN controller is done offline using VRFT by employing a Nonlinear Model Predictive Control (NMPC) scheme, and in the second tuning phase only the process sign is needed to back-propagate the error in order to adaptively correct the NN parameters/weights. The proposed tuning has as a main consequence the feedback linearization of the CS and represents another contribution of the paper.

After the low level feedback CS is fixed, a model-free Iterative Learning Control (MFILC) approach is proposed at the intermediate level to optimize the reference input/controlled output behaviors called primitive pairs, or simply primitives. The learning is performed under a criterion that balances execution accuracy and execution effort as well.

These reference input/controlled output pairs are stored into the memory under the form of a database (DB). The convergence of the learning scheme can be tackled based on the enforced CS description via the NN-VRFT design.

When a new task to be executed is dictated by a high level planner, the already optimized primitives are adaptively merged at the high level based on a similarity and selection criterion rather than by using a model of the CS to explicitly compute offline optimal solutions to the new task. This approach adaptively computes the reference inputs to address a diversity of requirements that can occur in the execution of the new task. For example, at different execution stages, different speeds may be required, or different efforts, or even different precisions. The resulting model-free adaptive control based on merging optimized primitives is another contribution of this paper.

Several approaches to primitive-based learning in the theoretical framework of CSs are presented in the literature. As shown in (Radac and Precup, 2015a), these approaches are organized in three categories: time-scale transformation approaches, temporal concatenation of primitive-based approaches, and time-based decomposition approaches. These categories are briefly discussed as follows.

Recent time-scale transformation approaches are presented in (Kawamura and Sakagami, 2002; Ijspeert et al., 2002). An Iterative Learning Control (ILC)-based approach is suggested in (Kawamura and Sakagami, 2002) to improve the maneuvers of an underwater robotic manipulator. A demonstration by learning approach is proposed in (Ijspeert et al., 2002). Each motion primitive is encoded in (Ijspeert et al., 2002) through the same nonlinear dynamic equations called attractor dynamics, which are invariant with respect to time scale, initial conditions and execution time. The motion equations are parameterized by Gaussian kernels and the parameters of each primitive are learned independently.

Temporal concatenations of primitive-based approaches are suggested in (Hoelzle et al., 2011; Schölling et al., 2011; Grymin et al., 2014). The feasibility of primitive motion tasks for Unmanned Aerial Vehicles is discussed in (Schölling et al., 2011). The Fourier series decomposition is applied to obtain the motion primitives needed in choreographic motion, and the temporal concatenation of primitives is proposed. The concept of library of motion primitives is suggested in (Hoelzle et al., 2011). A pair of input/output trajectories is available for each primitive, and these trajectories are learned by ILC. The temporal concatenation of primitives generates complex trajectories, and the Linear Time-Invariant (LTI) systems framework analysis is inserted in a bumpless transfer mechanism between primitives. An A^* search algorithm for the optimal temporal concatenation of primitives for mobile robot obstacle avoidance is proposed in (Grymin et al., 2014). The primitive execution control is optimally designed via Linear Matrix Inequalities in a hybrid Linear Time-Variant systems framework.

A time-based decomposition approach is given in (Wang and Zou, 2014). The primitives are B-spline functions, considered as elements of the library of primitives. The real-time planning of trajectories is performed using the learned B-spline primitives and combining them in the LTI systems framework.

None of the above primitive-based approaches rely on the model-free design of both the low level feedback controller and the high level ILC controller, which is different from our novel model-free approach.

VRFT (Previdi et al., 2004; Campi and Savaresi, 2006; Esparza et al., 2011; Formentin et al., 2013) belongs to data-driven or data-based model-free approaches employed in feedback controller tuning. Some of them are iterative experiment-based such as Iterative Feedback Tuning (IFT) (Hjalmarsson, 2002; Sjöberg et al., 2009), iterative Correlation-based Tuning (Mišković et al., 2007), Simultaneous Perturbation Stochastic Approximation (SPSA) (Spall and Cristion, 1998). Other model-free adaptive approaches include