Chapter IX
Monocular Vision System that Learns with Approximation Spaces

James F. Peters
University of Manitoba, Canada

Maciej Borkowski
University of Manitoba, Canada

Christopher Henry
University of Manitoba, Canada

Dan Lockery
University of Manitoba, Canada

ABSTRACT

This chapter introduces a monocular vision system that learns, with approximation spaces, to control the pan and tilt operations of a digital camera that is tracking a moving target. This monocular vision system has been designed to facilitate inspection by a line-crawling robot that moves along an electric power transmission line. The principal problem considered in this chapter is how to use various forms of reinforcement learning to control movements of a digital camera. Prior work on the solution to this problem was done by Chris Gaskett, using neural Q-learning starting in 1998, and more recently by Gaskett in 2002. However, recent experiments have revealed that both classical targets tracking as well as other forms of reinforcement learning control outperform Q-learning. This chapter considers various forms of the actor critic (AC) method to solve the camera movement control problem. Both the conventional AC method, as well as a modified AC method that has a built-in run-and-twiddle (RT) control strategy mechanism, is considered in this chapter. The RT mechanism, introduced by Oliver Selfridge in 1981, is an action control strategy, where an organism continues what it has been doing.
The problem considered in this chapter is how to guide action choices by an actor that is influenced by a critic governed by the evaluation of past actions. Specifically, one might ask how to measure the value of an action relative to what has been learned from experience (i.e., from previous patterns of behavior), and how to learn good policies for choosing rewarding actions. The solution to this problem stems from a rough-set approach to reinforcement learning by cooperating agents. It is an age-old adage that experience is a good teacher, and one learns from experience. This is at the heart of reinforcement learning, where estimates of the value of an action are based on past experience. Prior work on the reinforcement learning control was done by Chris Gaskett (2002), using neural Q-learning, starting in 1998. However, recent experiments have revealed that both classical target tracking, as well as other forms of reinforcement learning control, outperform Q-learning. Consideration of Gaskett’s particular version of Q-learning and neural networks as means of camera movement control is outside the scope of this chapter. This chapter considers various forms of the actor critic (AC) method to solve the camera movement control problem. AC methods have been studied extensively (see, e.g., Barto, Sutton, Anderson, 1983; Berenji, 2003; Bertsekas & Tsitsiklis, 1996; Konda & Tsitsiklis, 2000; Rosenstein, 2003; Sutton & Barto, 1998; Watkins & Dayan, 1992; Wawrzynski, 2005). The conventional actor critic method evaluates whether things have gotten better or worse than

**INTRODUCTION**

The problem considered in this chapter is how to guide action choices by an actor that is influenced by a critic governed by the evaluation of past actions. Specifically, one might ask how to measure the value of an action relative to what has been learned from experience (i.e., from previous patterns of behavior), and how to learn good policies for choosing rewarding actions. The solution to this problem stems from a rough-set approach to reinforcement learning by cooperating agents. It is an age-old adage that experience is a good teacher, and one learns from experience. This is at the heart of reinforcement learning, where estimates of the value of an action are based on past experience.

In reinforcement learning, the choice of an action is based on estimates of the value of a state and/or the value of an action in the current state. A swarm learns the best action to take in each state by maximizing a reward signal obtained from the environment. Two different forms of actor critic (AC) method are investigated in this chapter, namely, a conventional AC method and a form of AC method that includes an adaptive learning strategy, called run—and—twiddle (RT), played out in the context of remembered behavior patterns that accumulate in what are known as ethograms. An ethogram is a table of stored behavior patterns (i.e., vectors of measurements associated with behavior features) borrowed from ethology by Tinbergen (1963). Quantitative comparisons of past behavior patterns, with a template representing “normal” or desirable behavior, are carried out within the framework of an approximation space. Approximation spaces were introduced by Zdzislaw Pawlak (1981) during the early 1980s, elaborated by Orlowska (1982), Pawlak (1982), and generalized by Skowron and Stepaniuk (1995) and Stepaniuk (1998). The motivation for considering approximation spaces, as an aid to reinforcement learning, stems from the fact that it becomes possible to derive pattern-based action preferences (see, e.g., Peters & Henry 2005a, 2005b).

Prior work on the reinforcement learning control was done by Chris Gaskett (2002), using neural Q-learning, starting in 1998. However, recent experiments have revealed that both classical target tracking, as well as other forms of reinforcement learning control, outperform Q-learning. Consideration of Gaskett’s particular version of Q-learning and neural networks as means of camera movement control is outside the scope of this chapter. This chapter considers various forms of the actor critic (AC) method to solve the camera movement control problem. AC methods have been studied extensively (see, e.g., Barto, Sutton, Anderson, 1983; Berenji, 2003; Bertsekas & Tsitsiklis, 1996; Konda & Tsitsiklis, 2000; Rosenstein, 2003; Sutton & Barto, 1998; Watkins & Dayan, 1992; Wawrzynski, 2005). The conventional actor critic method evaluates whether things have gotten better or worse than...