ABSTRACT

Autonomous motion control of USVs, especially in complex marine conditions, is always a fundamental problem. Conventional methods consider the USV hydrodynamic model and the influences of environmental disturbance separately. However, due to the randomness of wind, wave and current, the accumulative error of each model can be large. To address this issue, this paper presents an end-to-end USV tracking control method via deep reinforcement learning, where a modern Reinforcement learning algorithm Actor-Critic is adopted. Given no prior knowledge of the dynamical system, the proposed method takes as input the information of environment (e.g., speed of wind and flow, etc.), ship and target trajectory, then produces the ship control signal (i.e., rudder angle and forward momentum) directly. We further propose a customized reward function to appraise the performance of ship agent. The presented simulation results demonstrate that this novel algorithm performs well in tracking tasks under complex marine conditions which is designed to change constantly.

KEY WORDS: unmanned surface vehicle, reinforcement learning, deep learning, trajectory tracking

INTRODUCTION

In recent years, unmanned surface vehicle (USV) is playing an increasingly import role in civil and military field. With the advantage of low-cost and multifunction, USV can easily be applied to many scenarios, like oceanographic measurement, ocean resource exploration, marine cruise and so on. In all kinds of tasks, trajectory tracking ability is essential for USVs and influences their performance. Traditionally, people build the dynamical model to calculate the hydrodynamic force (Fossen, Breivik, & Skjetne, 2003). Ivar-André F.Ihle applied passivity-based method to this problem(Ihle, Arcak, & Fossen, 2007). Marco Bibuli use Lyapunov-based guidance to guarantee the convergence of path-following(Bibuli, Bruzzone, Caccia, & Lapierre, 2009). And many other control methods are developed(Ashrafuon, Muske, & McNinch, 2010; Wang, Gao, Sun, & Zheng, 2018; Xiang, Yu, Lapierre, Zhang, & Zhang, 2018). When working in real world scenarios, the effect of the environment (i.e., wind, flow and wave) can not be neglected. A commonly used method is to build model for each force and much effort has been done (Fossen, 2012; Fossen & Lekkas, 2017; Fossen, Pettersen, & Galeazzi, 2015). However, it is notoriously challenging to calculate the effect of these disturbance precisely and there are always errors with the model, especially when it comes to wave, which is so complex that its real dynamics model can not be totally built yet. Moreover, the randomness of the environment will also increase the model errors. And when synthesizing all the models, the total error can even be larger, which may leads to the poor performance of USV in realistic environment.

On the other side, the artificial intelligence technique has made great progress in recent years. Mnih, Volodymyr and Koray Kavukcuoglu design a deep Q-network agent to play Atari 2600 games, whose performance is able to compete that of a professional human games tester(Mnih et al., 2013; Mnih et al., 2015). AlphaGo(Silver et al., 2016), a computer Go agent who defeat the best human player, can be regarded as a landmark, while its improved version AlphaGo Zero and AlphaZero (Silver, Hubert, et al., 2017; Silver, Schrittwieser, et al., 2017) can even be more capable. And the success of Libratus(Brown & Sandholm, 2017), the first AI to defeat top humans in heads-up no-limit Texas hold’em poker, also suggest that artificial intelligence agent is able to deal with imperfect-information tasks. All the progress above are based on deep reinforcement learning (DRL) technique. Reinforcement learning is about an agent interacting with the environment, learning an optimal policy, by trial and error, for sequential decision making problems in a wide range of fields in both natural and social sciences, and engineering(Sutton & Barto, 2018). Andrew Ng first successfully use a reinforcement learning algorithm to realize aerobatic helicopter flight(Abbeel, Coates, Quigley, & Ng, 2007). Jie Tan and Tingnan Zhang applied the RL-based actuator model to learning agile locomotion for quadruped robots(Tan et al., 2018). Pete Florence managed to train the robot arm quickly for a wide variety of previously unseen and potentially non-rigid objects(Florence, Manuelli, & Tedrake, 2018).

In this paper, we present an end-to-end USV motion control method via deep reinforcement learning where a modern reinforcement learning algorithm A3C(Mnih et al., 2016) is adopted. Instead of building the dynamical models for each part of drift forces, the proposed method construct multiple agents. These agents will be trained asynchronously.