論文 / 著書情報
Article / Book Information

題目(和文)	
Title(English)	Associative-memory based learning and concurrent deep reinforcement learning with application to autonomous aerial maneuver
著者(和文)	HuangPei-Hua
Author(English)	Pei-Hua Huang
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第10706号, 授与年月日:2017年12月31日, 学位の種別:課程博士, 審査員:長谷川 修,山村 雅幸,寺野 隆雄,青西 亨,石井 秀明
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第10706号, Conferred date:2017/12/31, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	論文要旨
Type(English)	Summary

（博士課程）
Doctoral Program

論 文 要 旨

THESIS SUMMARY

専攻： Department of	知能システム	専攻	申請学位（専攻分野）： Academic Degree Requested	博士 Doctor of	（工学） Engineering
学生氏名： Student's Name	Huang Pei-Hua		指導教員（主）： Academic Supervisor(main)	長谷川修	
			指導教員（副）： Academic Supervisor(sub)		

要旨（英文 800 語程度）
Thesis Summary（approx.800 English Words ）

Aerial robotic operations can be seen from recent year in many industries across the world, i.e., agriculture, construction inspection, and land surveying and mapping, today they can also be seen around our household. They have the ability to lift payload, travel through space with incredible speed. And imagining after catastrophic disturbance, instead of risking human life for the on-site investigation, single or multiple aerial robots automatically navigating through collapsed buildings to conduct search and rescue mission. They enter building, a few start mapping the internal structure of the building, and a few joint the fleet to search for survivors. They have not seen the environment before, nor have they seen the particular obstacles before, yet they are capable of self-localizing that they know where they are in the new environment, and manages to avoid obstacles efficiently along the way of searching for targets. This thesis discussed our research goal is to find the feasibility of using 3D simulations or real-world hardware to train a control policy for micro aerial vehicle (MAV) to perform various tasks of maneuver manipulation from scratch, the MAV will learn to navigate without any prior knowledge of the environment, using only sensory data as input. Achieving this would help us to discover a generalized model for navigation tasks, solve practical problems of model-free algorithms, and open up the possibility of transferring learned policy representation from a simulation setting to a physical hardware setting without any reinforced training.

We discussed approaches in our thesis of how to accomplish optimal control of MAV with learning based methods, by applying various model-free methods to achieve autonomous navigation from initially unknown environments for computationally constrained micro aerial vehicles (MAV), and discover efficient control policy representations with the focus on learning based methods, we seek to identify appropriate model-free approaches for both, the learning efficiency and (near-) optimal control policy. In this dissertation, we describe learning algorithms with formal performance guarantees which show that each of proposed methods can effectively solve control problems of maneuvering micro aerial vehicles — mainly address the problems of model-free learning methods for (i) Provide algorithms for improving data collection and data fitting challenges, improve the exploration. (ii) learn to perform maneuver tasks without any prior knowledge. We summarize the following structural outline in the respective chapters of the thesis:

Chapter 1. An introduction of an overall view for this thesis, this Chapter includes descriptions of the objectives and ultimate goal for this thesis, as well as the challenges that encountered during our works. Finally, summarize the contributions of this thesis.

Chapter 2. Following with Chapter 1 of an comprehensive introduction of the thesis, it introduces the background of core material covering the fields of hardware development of micro aerial vehicle (MAV), and its general dynamic models that most simulator used. We then cover the perception part of the MAVs, especially the self-localization concept for MAVs state estimation. Finally, provide relevant state-of-the-art works in these respective fields.

Chapter 3. We provide, a presentation of our first achievement on real world MAVs - with

the associative memory based method to self-organizing incremental neural network (SOIAM). In this first method – a model free supervised learning method – we highlight our modifications to adapt the associative memory concept to the self-organizing neural network (SOINN), in this chapter, we introduce the training and evaluating procedure of the quadcopter maneuvers in real world settings for hovering task. The results show that it is sufficient to learn a dynamics model from manual demonstrations.

Chapter 4. After we analyzed previous algorithms that described in Chapter 3, in order to resolve the problems raised in Chapter 3, we concentrate on our second method in concurrent deep reinforcement learning – the asynchronous trust region policy optimization (ATRPO) – In Chapter 4, we present both asynchronous and synchronous frameworks combined with reinforcement learning algorithm that resolve the exploration problem in both the SOIAM method and general model-free learning methods, we also provided empirical studies of continuous Mujoco simulation, eventually evaluated the results of empirical studies of the proposed method, that demonstrated improvement on the learning efficiency of tasks with different forms of parallel learning paradigm. Additionally, described how the proposed system can improve the issues of data acquisition and exploration in our first method that described in Chapter 3. Furthermore, the resulting design of the system will then enable us later to manipulate velocity control of the the simulated quadcopter for specific tasks.

Chapter 5. In this chapter, we test our previously proposed systems that are theoretically analyzed in a simulated quadcopter setting in real-world scenario, with tasks: hovering and inverted pole balancing. These experiments examined and evaluated different sensitivities of the proposed system according to various quadcopter task complexities and demonstrated a functioning, model-free concurrently learning based system for quadcopter application. Also the experimental results show the proposed frameworks outperforming the conventional methods by gaining faster convergence of learning a good control policy.

Chapter 6. In this last chapter comprehensively concluded this thesis with additional literature reviews that related to the achievements of this thesis, and discussed the feasibility of future works of combining SOIAM method and ATRPO method to improve the stability of learning, as well as some high level applications such as transferring from simulation setting into real world setting

.