Robotics and Autonomous Systems 100 (2018) 108–118

Contents lists available at ScienceDirect

Robotics and Autonomous Systems

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

A saliency-based reinforcement learning approach for a UAV to avoid flying obstacles

Zhaowei Ma, Chang Wang, Yifeng Niu *, Xiangke Wang, Lincheng Shen

College of Mechatronic Engineering and Automation, National University of Defense Technology, Changsha 410073, China

HIGHLIGHTS

- We propose a novel framework in which a UAV can learn to avoid flying obstacles.
- We propose an improved saliency detection method based on convolution neural networks.
- We propose an actor–critic reinforcement learning algorithm to control UAV in continuous spaces.

ARTICLE INFO

Article history:
Received 15 February 2017
Received in revised form 12 September 2017
Accepted 13 October 2017
Available online 14 November 2017

Keywords:
UAV
Flying obstacle avoidance
Convolution neural networks based saliency detection
Reinforcement learning

ABSTRACT

Obstacle avoidance is a necessary behavior to guarantee the safety of an unmanned aerial vehicle (UAV). However, it is a challenge for the UAV to detect and avoid high-speed flying obstacles such as other UAVs or birds. In this paper, we propose a generic framework that integrates an autonomous obstacle detection module and a reinforcement learning (RL) module to develop reactive obstacle avoidance behavior for a UAV. In the obstacle detection module, we design a saliency detection algorithm using deep convolution neural networks (CNNs) to extract monocular visual cues. The algorithm imitates human's visual detection system, and it can accurately estimate the location of obstacles in the field of view (FOV). The RL module uses an actor–critic structure that chooses the RBF neural network to approximate the value function and control policy in continuous state and action spaces. We have tested the effectiveness of the proposed learning framework in a semi-physical experiment. The results show that the proposed saliency detection algorithm performs better than state-of-the-art, and the RL algorithm can learn the avoidance behavior from the manual experiences.

1. Introduction

Unmanned aerial vehicles (UAVs) have been widely used in many domains such as aerial photography, search and rescue, anti-terrorist action, and agricultural surveillance. However, avoiding flying obstacles remains a big challenge for integrating UAVs into the national airspace system (NAS) with reliable and safe operations. Manual control may reduce the risk of UAV collisions under certain circumstances, but it requires substantial training and imposes significant cognitive loads on human operators, especially when a swarm of UAVs are controlled. In addition, it is difficult for the human operators to find obstacles at high altitude or in low visibility using the human visual system (HVS). Furthermore, both the UAV and the flying obstacles are moving in real-time and the dynamics of the flying obstacles are not always available. Therefore, it is necessary to develop an autonomous obstacle detection and avoidance method that can handle dynamic environments and learn from the human operators.

In the literature, methods combined with a low-level reactive layer have proved effective in solving the obstacle avoidance problem [1–6]. In the low-level reactive layer, the biological inspired behavior-based mode (BBM) [1] maps from perception to behavior directly, which is reactive to the changing environment and runs fast. The BBM mainly mimics how flying insects avoid the obstacle. However, these methods are typically based on specific environment models which are difficult to acquire [4–6]. In contrast, we take a model-free robot learning approach to develop adaptive obstacle avoidance behavior. Specifically, we choose the actor–critic (AC) reinforcement learning (RL) framework which is a machine learning technique based on trial-and-error mechanisms, thus improving the performance by getting feedback from the environment. It does not require knowledge of the environment and supports online learning.

Another challenge for avoiding obstacles for UAVs is to detect the obstacles. Much research have been done on automatic obstacles detection using computer vision methods. Among them,

* Corresponding author.
E-mail address: niuyifeng@nudt.edu.cn (Y. Niu).

https://doi.org/10.1016/j.robot.2017.10.009

optical flow is a bio-inspired approach that many flying insects use to sense the collision [7]. However, an embedded optical flow is limited in its resolution, providing only general guidance about obstacles. An estimated depth map using multiple images can provide obstacle distance information [8]. However, these algorithms require handcrafted features, which decreases the generalization ability of handling unknown environments. In this paper, we choose the saliency-based approach to detect the obstacles using a monocular forward-looking camera. This method is inspired by the human saliency system, which is a neural cognitive reaction controlled by human brains. And a lot of works have been done to develop this method [9–13]. Unlike their methods where mid-level filters are handcrafted, the saliency detection method based on deep convolution neural network (CNN) in our framework is automatically and jointly learned in a discriminative manner, because the deep networks mimic the functions of neocortex in the human brain as a hierarchy of filters and nonlinear operations, which improves the robustness of the algorithm.

The main contributions of this paper are as follows:

(1) We propose a novel framework in which a UAV can learn to avoid flying obstacles. This framework integrates an autonomous obstacle detection module and a reinforcement learning (RL) module to develop reactive obstacle avoidance behavior for a UAV. By using this framework, the UAVs can take the right obstacle avoidance action at the right time based on the environmental states.
(2) We propose an improved saliency detection method based on convolution neural networks. By using the deep neural networks, the obstacle detection algorithm has more generic application without handcrafted features. Furthermore, we apply the Kalman filter (KF) to predict the moving trajectory of flying obstacles.
(3) We propose an actor–critic reinforcement learning algorithm to control the UAV in continuous spaces. In this paper, we use RL to construct a mapping from the environmental states to the avoidance action. We choose the radial basis function (RBF) neural network in an actor–critic reinforcement learning module. In this algorithm, a reward function is specially constructed according to the obstacle state, which is used to guide and update the actor and critic the fitting model.

The remainder of this paper is organized as follows. In Section 2, we briefly review related work on the saliency detection methods and vision-based reinforcement learning methods. In Section 3, we introduce the UAV obstacle avoidance task and propose the saliency-based learning framework. A saliency detection algorithm based on CNN is proposed in Section 4. In Section 5, the reinforcement learning framework shows how to map the visual states to the reference angles based on the RBF neural network. Section 6 presents the offline validation experiments about this learning framework by using the simulated environment and a control hardware unit. The saliency detection algorithm is also validated using the databases and images from the simulating system. Finally, we conclude and discuss the future work in Section 7.

2. Related work

We give a brief overview about saliency detection methods and the vision-based RL method in this section.

The saliency detection method is inspired by the human vision system (HVS). Visual saliency shows the region that attracts human attention. The HVS is able to identify salient objects even in a complex scene by exploiting the inherent visual attention mechanisms. The first proposed saliency model computes feature maps with luminance, color, and orientation using a center–surround operator across different scales [14]. Then it performs normalization and summation to generate the saliency map. There are two kinds of methods to compute visual saliency: the bottom-up and the top-down methods [9–13]. Generally, the bottom-up models are based on a center–surround scheme, computing a master saliency map by various types of handcrafted low-level features such as color, intensity, and texture. The top-down methods need high-level knowledge to obtain the location of a salient object. However, handcrafted features are not robust enough for challenging scenes and to lose detailed, important and useful information.

Vision-based reinforcement learning (RL) methods have been applied in a variety of robotic tasks. For example, behaviors of other soccer robots are recognized for developing cooperative action policies [15–18]. Q-learning has been used for solving the simultaneous localization and mapping (SLAM) problem [19]. A deep unsupervised convolution network has been developed on the TORCS simulator [20]. In [21], neuro-evolution is combined with an unsupervised sensory pre-processor or compressor that is trained on images generated from the environment by the population of evolving recurrent neural network controllers. Reaching and pushing actions are learned to manipulate a box [22]. An intrinsically motivated RL approach has been proposed for active learning of object affordances and manipulation skills in continuous state and action spaces [23]. Similarly, a model-free actor–critic algorithm based on the deterministic policy gradient can handle continuous actions [24].

However, all these methods have limitations in solving the flying obstacle avoidance problem for UAVs, because the flying obstacles can be fast, small and blurred, which makes them hard to detect and track. Besides, learning control policies are also challenging in high-dimensional visual spaces and continuous action spaces. To address these challenges, we propose a novel method that integrates saliency-based object detection with the RL-based action selection. Our research is related to the saliency detection [9–13], but we have done this work using the CNN-based method.

3. Task description and the RL framework

3.1. Flying obstacle avoidance

In this paper, we discuss the problem of avoiding a single flying obstacle for UAV using a single camera. As shown in Fig. 1, the obstacle can be another aircraft or a bird. As the relative distance between the UAV and the obstacle changes, the size and location of the obstacle will change accordingly in the captured image, as shown in Fig. 1. The task is to detect the obstacle and observe its states, then select actions for the UAV to avoid the obstacle.

At the time step t, the UAV captures an image I_t with the size of $m = w \times h \times c$, where w is the width of the image, h is the height and c, is the channel. Denote by $s_t = \{u_t, v_t, O_t\}$ the state vector of the obstacle, where (u_t, v_t) is the center coordinates of the obstacle in I_t, and O_t is the number of pixels occupied by the obstacle. Then, we choose the reference turning angles $\begin{bmatrix}\psi_t & \phi_t\end{bmatrix}$as the action vector a_t for the UAV, where ψ_t is the yaw angle and ϕ_t is the roll angle. The goal state of avoiding the obstacle is described as $s_{t\_final} = \left\{u_{t\_final}, v_{t\_final}, 0\right\}$, where $u_{t\_final} \in \left\{-\frac{w}{2}, \frac{w}{2}\right\}$ and $v_{t\_final} \in \left[-\frac{h}{2}, \frac{h}{2}\right]$. The final state means that the flying obstacle is located at the margin of the image coordinate system.

3.2. Vision-based reinforcement learning framework

Based on the obstacle avoidance task description, we propose a saliency-based reinforcement learning framework as shown in