Contents lists available at ScienceDirect
Aerospace Science and Technology
www.elsevier.com/locate/aescte

Edge-based target detection for unmanned aerial vehicles using competitive Bird Swarm Algorithm

Xiaohua Wang, Yimin Deng, Haibin Duan

State Key Laboratory of Virtual Reality Technology and Systems, School of Automation Science and Electrical Engineering, Beihang University, Beijing 100083, PR China

A R T I C L E I N F O

Article history:
Received 1 June 2017
Received in revised form 26 April 2018
Accepted 30 April 2018
Available online xxxx

Keywords:
Unmanned aerial vehicle
Target detection
Edge potential function
Bird Swarm Algorithm

A B S T R A C T

Target detection for unmanned aerial vehicles is an important issue in autonomous formation flight. In this paper, a novel target detection approach for unmanned aerial vehicle formation is proposed based on edge matching. The windowed edge potential function is utilized to describe the attraction field for similar edges. Afterwards, the edge-based target detection problem can be formulated as an optimization problem. An improved version of the bird swarm algorithm, which is called competitive bird swarm algorithm, is proposed to find the location, rotation angle and scale of a given template on a specific image. A strategy named "disturbing the local optimum" is designed to help the original Bird Swarm Algorithm converge to the global optimal solution faster and more stably. Unmanned aerial vehicles moving in leader-follower pattern, which are called formation flight platforms, are used for our experiments. Images obtained by vision sensors embedded in the leaders are used to verify the effectiveness of the proposed method. The proposed algorithm is tested on both indoor and outdoor images to demonstrate the robustness. Comparative experiments with other state-of-the-art algorithms, including genetic algorithm, particle swarm optimization, artificial bee colony algorithm, pigeon-inspired optimization, and the basic bird swarm algorithm, are also conducted. The results prove the superiority and robustness of the proposed target detection algorithm.

1. Introduction

Unmanned Aerial Vehicles (UAVs) are important platforms in both civilian and industrial applications [1], for the advantages of zero casualties, good stealth performance, short operational preparation time, and relatively low life-cycle cost [2]. Multi-UAVs moving in formation have attracted much attention [3–8]. With visual sensors becoming more and more advanced, vision based formation has aroused much interest [9–12]. In leader-follower formation, the leader can ascertain the direction and distance of the following UAVs with information obtained by visual sensors. Therefore, target detection is an important task in vision based autonomous formation flight of UAVs. The aim of this study is to design an adaptive, efficient and robust target detection algorithm that can be applied to vision based UAV formation.

A wide variety of strategies have been established to deal with the target detection problem for UAVs in recent years [13]. A real-time detection algorithm for moving target from UAVs is proposed in [14]. A template matching approach is used in [15] to detect and track the run-way in image sequences. A target detection algorithm based on a visual attention model is given in [16]. A novel bio-inspired model is proposed in [17] via improved artificial bee colony and visual attention. An image registration algorithm is proposed in [18] for moving target detection. In addition, edge or template matching based methods, such as Charmfer matching [19] and Hausdorff distance matching [20], have also been extensively used for target detection [21].

In this paper, an edge matching based method is proposed for UAV target detection, which is much simpler and more efficient compared with feature-based algorithms. The edge of an image is detected and utilized in a matching procedure, which searches for the image patch with the highest similarity to a given edge template. A fast edge detection method based on structured forests is proposed in [22], which can obtain real-time performance and achieve state-of-the-art edge detection results. The edge matching method proposed in [23] utilizes the Canny edge detector. However, the edge map of a high-resolution UAV image extracted by Canny, would be enriched with textures, which will aggravate the computational complexity of edge matching. Therefore, the outstanding edge detection method based on structured forests is utilized to extract edges in this paper. Afterwards, the windowed edge potential function (WEPF) [23] is used to make a description

E-mail address: hbduan@buaa.edu.cn (H. Duan).

https://doi.org/10.1016/j.ast.2018.04.047
1270-9638/© 2018 Elsevier Masson SAS. All rights reserved.

of image edge. In terms of matching results and computational complexity, the WEPF based description is superior to traditional edge matching methods [24].

The edge-based matching problem can be considered as an optimization problem which can be resolved with various methods [13,21,23,24]. There are many advantages of bio-inspired optimized algorithms, such as high robustness, good distributed computing mechanisms, and extensive feasibility [25]. Thus a newly proposed bio-inspired optimization algorithm, Bird Swarm Algorithm (BSA) [26], is used for edge matching in aerial images. BSA is a novel optimization algorithm inspired by the social behaviors and social interactions in bird swarms. The searching space of BSA can be very large because of the high resolution of aerial images. In addition, the background of the scenes can be cluttered and complex, which means that there are many local optima in the edge matching problem. Thus, the basic BSA may be trapped into local optima or converge slowly. A strategy called "disturbing the local optimum" is designed and integrated into the basic BSA to overcome these shortcomings of it. The proposed BAS with the "disturbing the local optimum" strategy is called Competitive Bird Swarm Algorithm (CBSA). The "disturbing the local optimum" in CBSA can automatically check whether CBSA is trapped into local optima and add some disturbance to the local optimal solution if necessary. Therefore, the exploration ability of it is enhanced and the diversity of the swarm is improved. Furthermore, premature converge can be avoided.

The proposed target detection method is tested on aerial images obtained by the visual sensor embedded in UAVs. Experimental results demonstrate that the proposed edge matching algorithm can deal with the target detection problem for UAVs effectively. Furthermore, the performance of CBSA is compared with that of BSA, which demonstrates that the proposed strategy can improve the searching and converging ability of CBSA. Moreover, comparative experiments on CBSA and several state-of-the-art bio-inspired algorithms, including the particle swarm optimization (PSO) [27], artificial bee colony (ABC) algorithm [21], Genetic Algorithm (GA) [28], and Pigeon-inspired Optimization (PIO) algorithm [13] are conducted to demonstrate the advantage of the CBSA approach.

The remainder of this paper is organized as follows. In Section 2, the edge detection approach and the principle of WEPF is introduced. In Section 3, the BSA algorithm is presented. The "disturbing the local optimum" strategy is proposed in Section 4, where the detailed implementation procedure of the proposed edge-based target detection algorithm is provided. In Section 5, a series of experimental results are given to demonstrate the effectiveness of the proposed approach, followed by the concluding remarks given in Section 6.

2. Edge detection and the WEPF model

2.1. Edge detection method using structured forests

Edges generally exhibit patterns of local structure, such as straight lines or T-junctions [29]. The problem of predicting local edge masks can be formulated as a structured learning framework applied to random decision forests [22,30].

1) *Structured random forests*

Structured labels are utilized to determine the splitting function at each branch in the tree. For a $d \times d$ image patch, the annotation of it can be either a segmentation mask $y \in Y = \mathbb{Z}^{d\times d}$ or a binary edge map $y' \in Y' = \{0, 1\}^{d\times d}$. Both representations are utilized in this approach. All the structured labels y at a given node are robustly mapped to a discrete set of labels $c \in C$, $C = \{1, \ldots, k\}$. Similar structured labels are assigned to the same discrete label. Standard information gain measures can be evaluated on the discrete space. A mapping from Y to an intermediate space Z is defined to measure similarity over Y and calculate information gain. m dimensions of Z are sampled to reduce dimensionality, and a reduced mapping $\Pi_\phi : Y \to Z$ is obtained. To further reduce the dimensionality, principal component analysis (PCA) [31] is utilized. Then a straightforward map from Z to C is utilized to obtain the discrete labels. PCA quantization can be used to obtain the discrete label set C. The top $\log_2(k)$ PCA dimensions can be used to quantize a discrete label c as the assignment of z. In this paper we set $m = 256$ and $k = 2$. Each forest predicts a patch of edge pixel labels that are aggregated across the image to compute the final edge map.

2) *Input features*

In this paper, 32×32 image patches are used to predict 16×16 structured segmentation tasks. Two types of features are used: pixel lookups and pairwise differences. $x \in \mathbb{R}^{32\times 32\times K}$ is the feature vector, where K is the number of channels. Three channels in the CIE-LUV color space together with the normalized gradient magnitude at the original scale and half resolution scale are used. Four orientation channels are derived from the gradient magnitude channels. Thus the input feature has 13 channels. Each channel is blurred with a radius 2 triangle filter and down sampled by a factor of 2. A large triangle blur is used on each channel (8 pixel radius), and each channel is down sampled to a resolution of 5×5. Then candidate pairs are sampled and pairwise differences are computed. Thus the total dimension of a candidate feature is 7228.

3) *Mapping function*

A mapping $\Pi_\phi : Y \to Z$ is defined to train decision trees. $y(j)$ for $1 \le j \le 256$ denote the jth pixel of mask y. $z = \prod(y)$ is a large binary vector that encodes $[y(j_1) = y(j_2)]$ for each unique pair $j_1 \ne j_2$.

4) *Ensemble model*

The outputs of multiple trees in the random forests are combined to achieve robust results. The corresponding edge map y' is stored at each leaf node together with the learned mask y. Multiple overlapping edge maps $y' \in Y'$ can be averaged to obtain a soft edge response.

5) *Multiscale detection and edge sharpening*

The structured edge detector can be implemented on a multiscale version to enhance the performance of it. Three versions of resolution (1/2, 1, and 2) are computed and the results of the three edge maps are averaged after resizing to the original image dimensions. The edge maps can be sharpened optionally using local image color and depth values, with which the edge maps are better aligned to the image data. Additionally, the edge values vary from 0 to 1, which is continuous-valued. However, binary-valued edge information should be imported to the EPF computation. Thus, edge values lower than 1/3 of the maximum value in the corresponding edge image are set as 0, while others are set as 1 in this paper.

2.2. The principle of EPF

The image edges are considered as charged elements in EPF, which can generate an attraction field over object with similar edges. The concept of EPF is derived from the physics of electricity, simulating the electric potential generated by the electrostatic field. It is utilized in this paper to model the potential generated by edge structures of images. An edge template of a particular target is attracted by a set of equivalent charged edge points, which maximize the potential in EPF.

In the electricity, a set of point charges in a homogeneous background Q_i generates a potential, whose intensity can be calculated as follows.

$$v(\vec{r}) = \frac{1}{4\pi\varepsilon} \sum_i \frac{Q_i}{|\vec{r} - \vec{r}_i|} \tag{1}$$