



吉林大学学报(工学版)

Journal of Jilin University(Engineering and Technology Edition)

ISSN 1671-5497,CN 22-1341/T

《吉林大学学报(工学版)》网络首发论文

题目: 多无人机地面移动目标搜寻与定位
作者: 徐卓君, 王耀祥, 黄兴, 彭程
DOI: 10.13229/j.cnki.jdxbgxb20220425
收稿日期: 2022-4-18
网络首发日期: 2022-07-05
引用格式: 徐卓君, 王耀祥, 黄兴, 彭程. 多无人机地面移动目标搜寻与定位[J/OL]. 吉林大学学报(工学版). <https://doi.org/10.13229/j.cnki.jdxbgxb20220425>



网络首发: 在编辑部工作流程中, 稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定, 且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式(包括网络呈现版式)排版后的稿件, 可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定; 学术研究成果具有创新性、科学性和先进性, 符合编辑部对刊文的录用要求, 不存在学术不端行为及其他侵权行为; 稿件内容应基本符合国家有关书刊编辑、出版的技术标准, 正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性, 录用定稿一经发布, 不得修改论文题目、作者、机构名称和学术内容, 只可基于编辑规范进行少量文字的修改。

出版确认: 纸质期刊编辑部通过与《中国学术期刊(光盘版)》电子杂志社有限公司签约, 在《中国学术期刊(网络版)》出版传播平台上创办与纸质期刊内容一致的网络版, 以单篇或整期出版形式, 在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊(网络版)》是国家新闻出版广电总局批准的网络连续型出版物(ISSN 2096-4188, CN 11-6037/Z), 所以签约期刊的网络版上网络首发论文视为正式出版。

多无人机地面移动目标搜寻与定位

徐卓君¹, 王耀祥¹, 黄兴¹, 彭程²

(1. 吉林大学 通信工程学院 长春 130022; 2. 中国科学院 长春光学精密机械与物理研究所, 长春 130033)

摘要：本文针对多无人机地面移动目标跟踪任务中的目标搜寻与定位问题，为了提高多机搜寻的效率，降低实际搜寻过程中操作和通信延时带来的误差影响，提高定位精度：首先通过地面站为单一无人机手动搜寻框选目标并提取目标共享信息；然后他机根据共享信息利用 YOLOv3 自动检测并定位目标；最后设计了基于判别相关滤波器的延时误差补偿算法，对选定过程中的坐标误差进行估计与补偿；实验结果表明，上述方法有效提高了多机目标搜寻与定位的效率和精度。

关键词：模式识别与智能系统;多无人机;目标搜寻;误差补偿

中图分类号：TP242 **文献标志码：**A

DOI：10.13229/j.cnki.jdxbgxb20220425

Ground moving target search and location with multi-UAVs

XU Zhuojun¹ WANG Yao-xiang¹ HUANG Xing¹ PENG Cheng²

(1. College of Communication Engineering, Jilin University, Changchun, 130022; 2. Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033)

Abstract: This paper aims at the problem of target search and positioning in the multi-UAV ground mobile target tracking task. In order to improve the efficiency of multi-machine search, reduce the error impact caused by operation and communication delay, and improve the positioning accuracy in the actual search process: the man-machine manually searches for the frame to select the target and extracts the shared information of the target; then the other machine uses YOLOv3 to automatically detect and locate the target according to the shared information; finally, a KCF-based delay error compensation algorithm is designed to estimate the coordinate error in the selection process. Compensation; Experimental results show that the above method effectively improves the efficiency and accuracy of multi-machine search and positioning.

Key words: Pattern recognition and intelligent system; Multi-UAV; Target search; Error compensation

收稿日期：2022-4-18.

基金项目：吉林省科技发展计划项目 (20200201294JC); 吉林省科技发展计划项目 (20150204074GX).

作者简介：徐卓君 (1983-), 女, 副教授, 博士. 研究方向: 模式识别方法研究. E-mail: xuzhuojun@jlu.edu.cn.

0 引言

近年来随着科学技术的快速发展,无人机的软硬件水平不断提高,在军用领域和民用领域都得到了广泛应用。在无人机的诸多应用中,目标追踪作为一项重要的子任务,在如城市反恐、地面打击和灾后搜救等应用中发挥着重要作用,也因此愈发受到研究人员的关注。而在实际应用中,单架无人机视野范围有限,任务环境复杂,易出现目标遮挡、丢失等干扰,因此多无人机目标跟踪自然就应运而生。

多无人机目标跟踪的首要工作是根据任务要求快速搜寻待跟踪目标并获取其位置。目前多无人机目标跟踪的研究中对目标的搜寻探究较少,多数是在已知待跟踪目标的前提下进行研究,如文献[1]中关于多无人机协同目标最优观测与跟踪的研究,在文中明确了目标已知的研究前提;文献[2]中采用受限粒子滤波对目标进行对峙跟踪的研究中,也是默认待跟踪目标已经知晓。而在专门关于多无人机目标搜寻的研究中,所用方法大多耗时较长,如文献[3]中采用的强化学习方法来进行多无人机目标搜寻,虽然探究了强化学习方法在改领域的可用性,但是耗时较长,且是泛类搜索,无法准确搜索到待跟踪目标。文献[4]中采用了基于滚动时域控制的协同搜索决策方法,通过设计基于 MTPM 和 DPM 的寻优指标,有效降低了对目标的虚警率和漏检率,实现了对任务区域的可控回访,但是耗时较长,不适合在任务初始阶段对目标进行快速搜寻。

目标定位方面,根据所需的坐标类型获取方式也有所不同。目前在多机跟踪任务中需要的目标初始坐标类型主要有两种,分别为像素坐标信息和世界坐标信息;在多机跟踪任务中,跟踪算法需要目标的初始像素坐标作为参考信息,而控制算法则需要目标的世界坐标信息。若需要目标的像素坐标信息,目前一般通过地面站利用 OpenCv 等软件进行框选即可获得;若需要目标的世界坐标信息,则通过多机联合定位或者单机自主定位去获得。多机联合定位如西北工业大学的屈耀红等人提出的利用多无人机和目标之间的距离以及无人机坐标信息来估算目

标位置的协同定位算法[5],该算法通过利用一阶泰勒展开,将非线性距离量测方程分化为量测距离信息、无人机站址误差和测距传感器的测距误差信息方程这三个部分,然后求解方程来获取目标位置,并利用在线枚举法寻求 Geometric Dilution Precision(GDOP)值映射的最优编队队形来优化坐标精度。而在单机自主定位方案中,文献[6]中 Joshua D.Redding 提出利用目标的像素坐标和无人机姿态及相机姿态角来进行目标定位的方案。文献[7]提出了一种基于共线构像原理实现简单、成本低的目标定位方法,但在地势起伏过大的地方误差会较大。文献[8]提出了一种基于滤波定位和视觉定位的混合定位方法,先通过滤波定位粗略估计目标位置,再通过视觉定位提高定位精度。文献[9]通过将 2D 激光测距仪和单目相机进行结合实现了目标选定后自动获取其世界坐标系坐标的功能。

国内外研究在多无人机地面目标搜寻及定位方面都各有建树,但尚存在以下不足:1)多数跟踪研究默认目标已确定,而目前关于目标搜寻的方法耗时长且无法搜索特定目标,不适用于多机跟踪任务的初始目标搜寻与定位;2)目前在多机跟踪任务中目标初始像素坐标信息是不可或缺的,但目前的研究中尚未关注到在实际跟踪过程中由于目标处于非静止状态,像素坐标获取过程的操作延时会给坐标带来误差;此外地面站与无人机之间的通信误差在各项定位研究中也考虑较少。

针对上述问题,本文通过地面站为机群中的某架无人机手动搜寻定位待跟踪目标,并提取目标的相关信息共享给他机,他机利用目标检测算法搜寻定位待跟踪目标,实现了多机对特定目标的快速搜寻与定位;针对当前研究中考虑较少的像素坐标获取过程中的操作延时以及地面站与无人机之间可能产生的通信延时问题,设计了基于判别相关滤波器(kernelized correlation filters,后简称 KCF)的延时误差补偿算法,利用 KCF 在不同帧之间定位目标坐标,结合无人机机载相机视频帧率获取相邻帧时间差,从而估计目标速度,最后结合无人机与地面站

的通信时间差来估计并补偿误差量，提高了实际应用中的目标定位精度。

1 地面目标搜寻及定位

在跟踪任务中，首先需要搜寻并定位待跟踪目标，利用地面站为无人机手动搜寻待跟踪目标能够更好的匹配任务需求，由地面人员可根据无人机的回传图像结合任务要求手动快速搜寻并选定待跟踪目标。在多无人机跟踪任务中，可选择通过地面人员为单台无人机手动选定目标（算法流程如表 1 所示），通过将解析出目标的全局坐标等信息共享给其他无人机（算法流程如表 2 所示），让他机自行检测识别出待跟踪目标，来完成多无人机的目标搜寻及定位。

表 1 地面节点获取伪代码

Table.1 Pseudocode of ground nodes

算法 1 地面节点
1: 获取无人机 x 的回传图像
2: while 未准确获取目标:
3: 根据任务要求手动搜寻、框选目标
4: 人工确认搜寻结果
5: end while
6: 提取目标共享信息
7: while 未收到他机确认信息:
8: 发送共享信息到他机
9: end while
10: 回传目标坐标到无人机 x 并开始跟踪任务

表 2 共享节点设定伪代码

Table.2 Pseudocode of shared nodes

算法 2 共享节点
1: while 未收到目标共享信息
2: 接收地面站目标共享信息
3: 确认信息完整性
4: end while
5: 发送接收确认信息
6: 根据共享信息使用 YOLOv3 检测并获取目标坐标

1.1 目标共享信息

无人机 x 的图像传回地面站后，地面人员根据任务要求手动搜寻待跟踪目标将其手动框选，框选后使用 opencv 自动获得目标的像素坐标(u,v)。由于各无人机的相机视角不同，导致目标在不同无人机

的像素坐标系中的坐标也不同，因此不能直接共享目标的像素坐标。

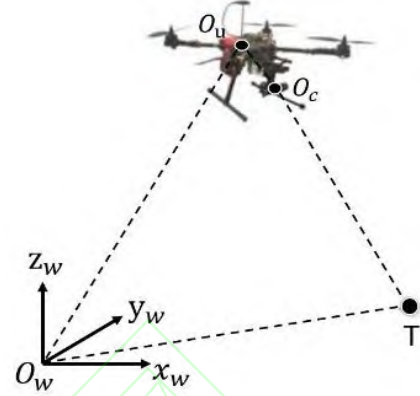


图 1 定位原理图

Fig.1 Positioning schematic

在与无人机相关的各坐标系中，以起飞点为原点的世界坐标系只与无人机起飞点位置有关，当各无人机从同一起飞点起飞时，各机所构建的世界坐标系是相同的，因此在以同一原点构建的世界坐标系中目标的坐标是相同的，不受无人机视角的影响。在世界坐标系中，如图 1 所示，目标的世界坐标表示如（1）所示：

$$\begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} = L_{uav} + L_t + RL_{uc} \quad (1)$$

其中， L_{uav} 表示无人机在世界坐标系的位置，通过无人机的 GPS 坐标获得； L_t 表示相机坐标系内的目标坐标在世界坐标系的投影； R 表示相机坐标系与世界坐标系之间的旋转矩阵，通过无人机飞行时的欧拉角获得， R 的具体表达式如下（2）所示：

$$R = \begin{bmatrix} R_1 & R_4 & R_7 \\ R_2 & R_5 & R_8 \\ R_3 & R_6 & R_9 \end{bmatrix} \quad (2)$$

$$= \begin{bmatrix} \cos\theta\cos\psi & -\cos\theta\sin\psi + \sin\phi\sin\theta\cos\psi & \sin\phi\sin\psi + \cos\phi\sin\theta\cos\psi \\ \cos\theta\sin\psi & \cos\theta\cos\psi + \sin\phi\sin\theta\sin\psi & -\sin\phi\cos\psi + \cos\phi\sin\theta\sin\psi \\ -\sin\theta & \sin\phi\cos\theta & \cos\phi\cos\theta \end{bmatrix}$$

其中， θ 、 ψ 、 ϕ 分别为俯仰角、偏航角和翻滚角。

L_{uc} 表示相机中心与无人机中心的位移，该值固定且可通过事先标定获得。故想获得目标的世界坐标，须解出相机坐标系内的目标坐标在世界坐标系的投影 L_t 。

首先推导坐标从像素坐标系到图像坐标系的转换过程：如图 2 所示，像素坐标系以 CCD 图像平面左上角的顶点为坐标原点，坐标轴分别平行于图像坐标系的两条坐标轴， (u,v) 表示在该坐标系中的坐标，其物理意义是构成目标的像素点在成像画面的像素表达平面中的坐标，用 (x,y) 来表示图像坐标系中的坐标值，其物理意义是用物理单位来表示构成目标的像素在成像画面中的位置。

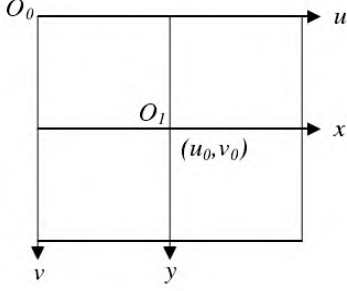


图 2 像素坐标系与图像坐标系关系图

Fig.2 Pixel coordinate system and image coordinate system relationship diagram

则坐标在像素坐标系和图像坐标系之间进行转换的过程如下：

$$u = \frac{x}{dx} + u_0 \quad (3)$$

$$v = \frac{y}{dy} + v_0 \quad (4)$$

其中 dx, dy 表示一个像素单位在坐标轴上的实际物理尺寸，将式(3)、式(4)转换成矩阵表达，便于后续计算，转换结果如下式所示：

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} 1/dx & 0 & \mu_0 \\ 0 & 1/dy & \nu_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (5)$$

接下来推导从图像坐标系到相机坐标系的转换：相机坐标系的原点为相机光心，而其 X 轴、 Y 轴分别与相机成像画面的 X 轴、 Y 轴平行，而 Z 轴则为相机的光轴，用 (x_c, y_c) 表示相机坐标系下的坐标值。各坐标系之间的关系如图 3 所示，坐标在图像坐标系和相机坐标系之间的转换关系如下：

$$x = f \times \frac{x_c}{z_c} \quad (6)$$

$$y = f \times \frac{y_c}{z_c} \quad (7)$$

其中 f 为相机焦距，将式(5)、式(6)、式(7)转换成矩阵表达：

$$\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = \begin{bmatrix} (u-u_0)/f_x \\ (v-v_0)/f_y \\ 1 \end{bmatrix} z_c \quad (8)$$

其中， $f_x = f/dx$ ， $f_y = f/dy$ 。计算目标的相机系坐标在世界坐标系下的投影：

$$\begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix} = R \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = \begin{bmatrix} R_1 & R_4 & R_7 \\ R_2 & R_5 & R_8 \\ R_3 & R_6 & R_9 \end{bmatrix} \begin{bmatrix} (u-u_0)/f_x \\ (v-v_0)/f_y \\ 1 \end{bmatrix} z_c \quad (9)$$

上式中 z_t 为无人机高度 h ，通过无人机自身传感器获得；上式可进一步写为：

$$\begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix} = R \begin{bmatrix} \frac{u-u_0}{f_x} \\ \frac{v-v_0}{f_y} \\ 1 \end{bmatrix} z_c = \begin{bmatrix} R_1 \frac{u-u_0}{f_x} + R_4 \frac{v-v_0}{f_y} + R_7 \\ R_2 \frac{u-u_0}{f_x} + R_5 \frac{v-v_0}{f_y} + R_8 \\ R_3 \frac{u-u_0}{f_x} + R_6 \frac{v-v_0}{f_y} + R_9 \end{bmatrix} z_c \quad (10)$$

由以上可得：

$$L_t = \begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix} = \begin{bmatrix} (R_1(u-u_0)/f_x + R_4(v-v_0)/f_y + R_7)z_c \\ (R_2(u-u_0)/f_x + R_5(v-v_0)/f_y + R_8)z_c \\ h \end{bmatrix} \quad (11)$$

其中：

$$z_c = h / (R_3(u-u_0)/f_x + R_6(v-v_0)/f_y + R_9) \quad (12)$$

然后由式(1)、式(2)及式(11)即可得目标得世界坐标。

最后将目标的全局坐标、目标类别和目标特征信息一齐通过地面站共享给其他无人机。

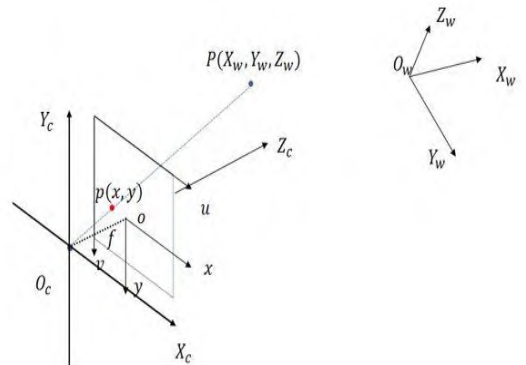


图 3 各坐标系之间关系

Fig.3 Relationship between coordinate systems

1.2 目标检测

他机接收到全局坐标后，根据自身位姿信息反解出像素坐标，并以该坐标为中心，按照相机画幅的一定比例裁剪获取的视频图片，这样做是为了提高算法速度来快速检测目标。检测算法采用了YOLOv3 算法，该算法可以自动检测出相应类别的目标并计算给出目标尺度框坐标，从而完成目标搜寻和定位。

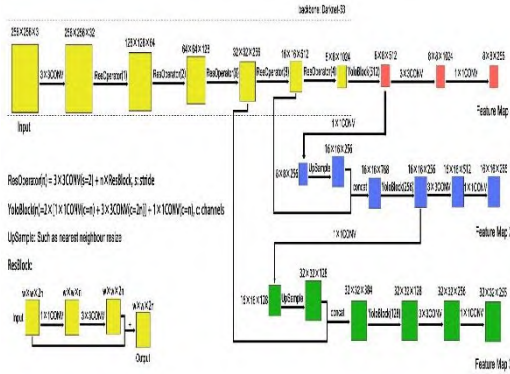


图4 YOLOv3 网络结构

Fig.4 Network structure of YOLOv3

YOLOv3 的网络结构如图 4 所示，相较于前一个版本的 YOLO 算法，YOLOv3 使用 Darknet-53 来代替之前的 Darknet-19 架构。Darknet-53 用于输入图像的特征提取，并通过增加卷积层数和引入 Resnet 中的跨层加和操作来提高精度。Darknet-53 的结构图如图 5 所示。

	Type	Filters	Size	Output
1x	Convolutional	32	3 × 3	256 × 256
	Convolutional	64	3 × 3 / 2	128 × 128
	Convolutional	32	1 × 1	
	Convolutional	64	3 × 3	
	Residual			128 × 128
2x	Convolutional	128	3 × 3 / 2	64 × 64
	Convolutional	64	1 × 1	
	Convolutional	128	3 × 3	
	Residual			64 × 64
	Convolutional	256	3 × 3 / 2	32 × 32
8x	Convolutional	128	1 × 1	
	Convolutional	256	3 × 3	
	Residual			32 × 32
	Convolutional	512	3 × 3 / 2	16 × 16
	Convolutional	256	1 × 1	
8x	Convolutional	512	3 × 3	
	Residual			16 × 16
	Convolutional	1024	3 × 3 / 2	8 × 8
	Convolutional	512	1 × 1	
	Convolutional	1024	3 × 3	
4x	Residual			8 × 8
	Avgpool		Global	
	Connected		1000	
	Softmax			

图5 Darknet-53 结构

Fig.5 Structure of Darknet-53

在 YOLOv3 的结构图中只包含卷积层，特征图

的尺寸大小由卷积步长来控制，因此对输入图片大小没有特殊限制。而从结构图中可以看到，YOLOv3 会输出三个不同尺度的特征图，这是因为 YOLOv3 采用了特征金字塔的方法，利用尺寸较小的特征图来检测大物体，利用大尺寸特征图来检测小物体，也正因此 YOLOv3 获得了更好的小目标检测性能。

输入图片首先经过 Darknet-53 处理，然后 YOLOblock 会对处理后的图片进一步操作，经 YOLOblock 生成后的特征图经过 3*3 的卷积层和 1*1 的卷积层后即可得到特征图 1；若经过 1*1 的卷积层和采样层，再与 Darknet-53 的中间层进行拼接，即可得到特征图 2；而特征图 3 是经由同样的循环所得。这三个不同尺度的特征图均包含了检测框的位置、检测置信度和检测类别，这些信息汇在一起就是 YOLOv3 的检测结果。

2 延时误差补偿

2.1 KCF 算法

KCF 算法源于信号处理领域，用于描述两个信号之间的相关性^[13]；而应用至目标跟踪领域后，则用以描述输入图像 I 与目标图像模板 T 之间的相关性，描述结果表述如下：

$$p = I \otimes T \quad (13)$$

其中 p 为输入图像的响应，输入图像 I 与目标图像模板之间 H 的相关性越高，则响应值 p 越大。为了提高计算速度，通常将式(13)转换到傅利叶域进行计算如下：

$$F(p) = F(I \otimes T) = F(I) \cdot F(T)^* \quad (14)$$

具体到 KCF 算法中，会在初始帧根据框选出的目标区域来训练目标检测器以获取目标图像模板。在框选区域内提取目标的 HOG 特征作为正样本，即 $P_x = [x_1, x_2, \dots, x_n]$ ，利用循环矩阵 C 将目标周围区域构造为负样本 $N_x = C \cdot P_x$ ，其中循环矩阵 C 如下式所示：

$$C = \begin{bmatrix} 0 & 0 & 0 & \dots & 1 \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix} \quad (15)$$

最后得到样本特征矩阵 X 如下：

$$\mathbf{X} = \begin{bmatrix} P_x \\ N_x \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \\ x_n & x_1 & \cdots & x_{n-1} \\ \vdots & \vdots & \cdots & \vdots \\ x_2 & x_3 & \cdots & x_1 \end{bmatrix} \quad (16)$$

得到正负样本后，算法首先利用高斯核函数将线性的特征空间映射到高维空间，映射操作如下：

$$\omega = \sum_i \alpha_i \varphi(x_i) \quad (17)$$

其中 ω 为映射后的特征。KCF 算法利用这些样本进行快速训练。

而在新一帧，KCF 算法会根据初始帧的目标位置划分目标搜索区域，并提取搜索区域的 HOG 特征作为输入 f ，并将其进行快速傅利叶变换，然后根据核相关滤波参数 α 计算各候选区域的回归值响应图 p ，如下式所示：

$$p = \mathcal{F}^{-1}(\alpha \odot k(f, x)) \quad (18)$$

响应图中最大响应值所对应的位置即为当前帧目标坐标。得到当前帧坐标后，算法在当前帧目标位置处划分搜索区域，提取搜索区域的 HOG 特征并对其做快速傅里叶变换，计算核相关滤波参数，最后对目标模板和核相关滤波参数进行更新如下：

$$x_{curr} = (1 - \gamma)x + \gamma f' \quad (19)$$

$$\alpha_{curr} = (1 - \gamma)\alpha + \gamma \alpha' \quad (20)$$

其中， f' 为得到当前帧坐标后，算法在当前帧目标位置处划分搜索区域的 HOG 特征， α' 为计算核相关滤波参数。在后续帧中 KCF 算法将重复以上步骤，实现在不同帧定位目标并获取其像素坐标。

2.2 坐标误差估计及补偿

令地面站在接收到的首帧无人机图像中手动框选得到的目标像素坐标为 (u, v) ，KCF 算法根据该坐标对框选后的目标进行跟踪，第 N 个目标像素坐标记为 (u_N, v_N) ，设无人机机载相机的帧率为 M 帧/秒，则相邻帧时间差为 $1/M$ 秒，则像素坐标系两个坐标轴方向上目标的速度分别为：

$$\begin{aligned} x_v &= \frac{u_N - u}{N} \times M \\ y_v &= \frac{v_N - v}{N} \times M \end{aligned} \quad (21)$$

为了提高式 (19) 中计算出的速度精度，在选定过程中，无人机会传回共 K 张图像，相邻图像间隔十帧。在第一张图像上完成目标框选后，KCF 会

在该张图片提取样本并进行训练，获得目标模板 h ，在之后的每一张图片上定位目标并计算目标的速度。最后对获得的 $k-1$ 次目标速度做加权均值处理，得到估计的目标速度为：

$$\begin{aligned} x_v &= \frac{\sum_i^{k-1} x_{vi} \times i}{\sum_i^{k-1} i} \\ y_v &= \frac{\sum_i^{k-1} y_{vi} \times i}{\sum_i^{k-1} i} \end{aligned} \quad (22)$$

设无人机从传回首张图像到接收目标坐标的两次通信时间差为 t ，补偿后待跟踪目标的坐标如下：

$$\begin{aligned} x &= u + x_v t \\ y &= v + y_v t \end{aligned} \quad (23)$$

3 实验分析与验证

3.1 多机目标搜寻实验

受飞行环境网络和通信设备限制，为规避通信对实验结果带来的干扰以及飞行画面传输不畅对框选操作的影响，本实验采用采集实际飞行数据进行离线实验。由于数据来源与实机飞行时相同，因此离线实验可以验证算法的正确性。实验平台硬件设备为：CPU R5-4600H，RAM 16GB。实验中，在实验平台同时运行两个跟踪节点，一个节点用于完成对目标的框选并将目标信息封装；另一节点在读取矩阵信息后运行 YOLOv3 算法检测目标。文中图 6 和图 7 为运行过程的实验系统界面截图，截图中左下角的图片为节点 1 视角，右上角图片为节点 2 视角。



图 6 信息共享前选定图

Fig.6 Selected map before information sharing

在图 6 中代表节点 1 的 tracking_uav1 已框选出

目标，并解算获得目标全局坐标和类别信息，但此时信息未共享给节点 2，因此代表节点 2 的 tracking_uav2 中目标人物未被自动检测、框选。

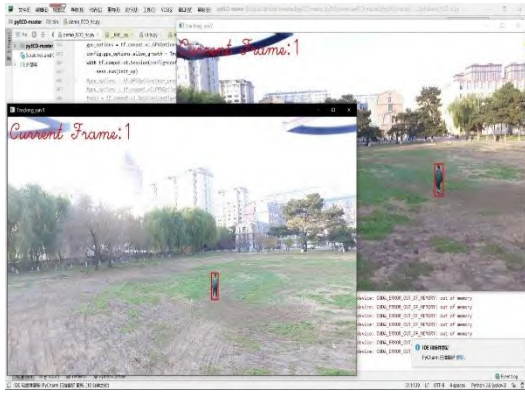


图 7 信息共享后选定图

Fig.7 Selected map after information sharing

在将节点 1 的信息共享给节点 2 后，如图 7 所示，代表节点 2 的 tracking_uav2 中已经自动检测出待跟踪目标并完成框选。可见 YOLOv3 算法能够依靠共享信息完成对目标的自动框选及定位，文章所设计的选定方案能够实现既定功能。在执行任务的机群数量较大时，能够有效提高目标选定的效率。

3.2 延时误差补偿算法验证实验

3.2.1 实验工况说明

根据无人机进行目标搜寻与定位的实际过程，实验中设计了以下几种不同的选定工况：1、无人机盘旋状态受扰动，目标匀速直线运动；2、无人机盘旋状态正常，目标变速变向运动；3、无人机盘旋状态受扰动、目标变速变向运动。

对于地面站返回目标坐标到无人机过程中可能面临的通信延时工况，受当前无人机通信设备影响，无法自由设置通信延时时间，难以模拟实际选定过程中可能遭遇的各种延时状况，在实验中实现精准的变量控制。因此采用离线实现以便于模拟不同的通信延时时长。

3.2.2 实验结果及分析

对于上文所述的三种选定工况，本文采用了 UAV123 数据集集中的 car3、car7 和 car2 测试集来分别模拟。而关于算法参数 k ：在实验进行前，通过不同人员和不同目标进行目标框选操作，统计得到该操作的平均耗时为 2-3 秒，而无人机机载相机的视频采集帧率为 30 帧每秒，因此无人机回传图像数量 k 被设置为 6，其中每张图片之间相隔十帧，图片传输总耗时为 $6 \times 10/30 = 2$ 秒，可以保证图片传输时间小于手动框选操作时间，避免了图片传输带来的额外延时。最后，通过比较算法输出的坐标与已标定

的目标真实坐标之间的误差大小来评判算法输出的坐标精度程度。

在实验工况 1 中对比了本文算法输出的坐标与未经补偿的、初始框选坐标之间的精度差别。如图 8 所示：横坐标为延时时间，由于图片传输时间小于等于手动操作时间，因此本文中延时自无人机图片传输完成后开始计。纵坐标为输出坐标与真实坐标在像素坐标系两个坐标轴方向上的坐标误差之和。手动框选得到的目标尺度框大小为：宽 15 像素、高 41 像素。图中红色曲线代表的是本文算法输出的坐标，蓝色曲线代表的是初始框选坐标，右上角为二者在不同延时情况下的平均误差。从曲线可以看到，随着延时时长的增加，本文算法输出的坐标误差始终维持在较低水平，而初始框选坐标在零延时下的误差已经超过 40 像素，并随着延时的增加，误差亦不断上升。

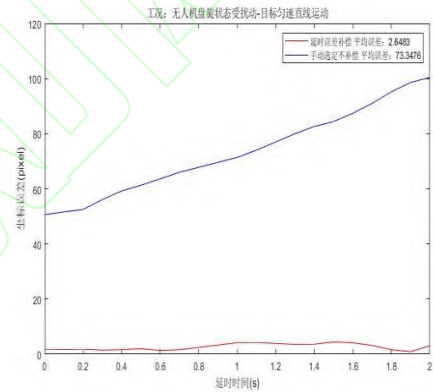


图 8 工况 1 误差曲线

Fig.8 Error curve of working condition 1

从平均误差看，初始框选坐标的误差高达 73.3476 像素，而本文算法的平均误差仅有 2.6483 像素，平均误差下降了 96.4%。

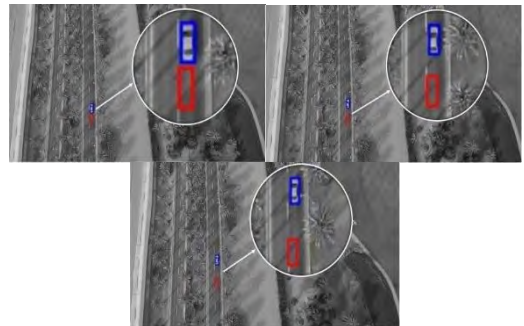


图 9 工况 1 不同延时结果对比实验图

Fig.9 Comparison experiment diagram of different delay results in condition 1

为了更直观的表现不同延时时长下本文算法输出的坐标与初始框选坐标的精度不同，提取了延时

时长为 0s、1s、2s 的实验图片，如图 9 所示，白色圆圈是为了便于观察做的放大效果，其中红色框代表初始框选的目标尺度框，蓝色框代表本文算法输出的目标；可以看到，在不同延时下本文输出的坐标框始终覆盖了目标，可为后续跟踪算法提供可靠的初始坐标参考信息；反观初始框选目标框在 0 延时情况下即已脱离目标，并随着延时时长增加与目标之间的距离也逐渐增大，显然若以初始框选坐标作为后续跟踪算法的初始参考坐标，必然会导致跟踪任务的失败。

在工况 2 中，手动框选的目标尺度框大小为：宽 44 像素、高 40 像素。实验结果如图 10 所示，可以看到随着延时时长的增加，本文算法输出的坐标误差和初始框选坐标误差都在上升，但在误差数值上，本文算法要低得多；从右上角的平均误差看，本文算法的误差值仅有 11.0238 像素，相较初始框选坐标的平均误差降低了 85.2%。

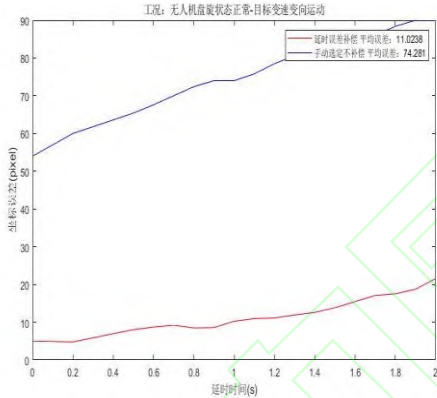


图 10 工况 2 误差曲线

Fig.10 Error curve of working condition 2

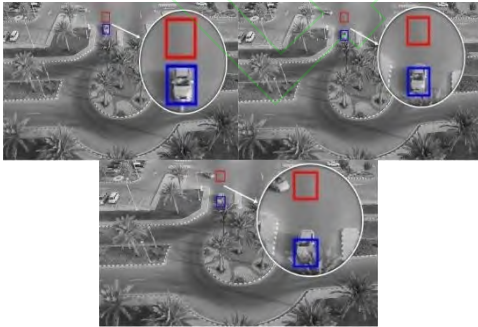


图 11 工况 2 不同延时结果对比实验图

Fig.11 Comparison experiment diagram of different delay results in condition 2

而从具体的实验图像看，依然提取延时时长 0s、1s、2s 的实验图片进行对比，如图 11 所示，可以看到红色框代表的初始框选目标框在 0 延时下即已丢失目标，随着延时时长的增加，与目标之间的距离

也进一步增大；反观本文算法，虽然延时时长增加后算法输出的目标框对目标的覆盖面积略有减少，但降低幅度较低，基本可以覆盖目标。

在工况 3 中，手动框选的目标尺度框大小为：宽 21 像素、高 36 像素。实验结果如图 12 所示，随着延时时长的增加，初始框选坐标误差不断上升，而本文算法输出的坐标误差在短暂上升后又趋于平稳。

而在平均误差上，本文算法的平均误差数值为 18.3857 像素，相比前两种工况都略有增加，但对比初始框选坐标的 115.5762 像素依然下降了 84.1%。

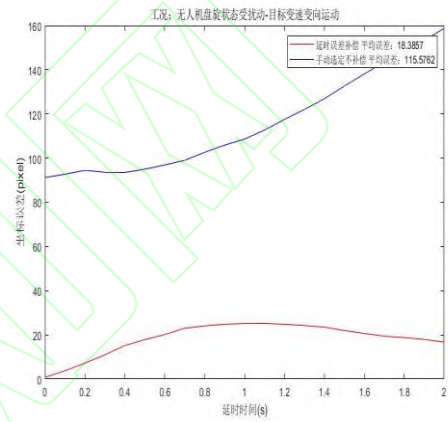


图 12 工况 3 误差曲线

Fig.12 Error curve of working condition 3

而从提取的不同延时下(延时时长分别为：0s、1s、2s) 的实验结果图来看，如图 13 所示，红色框代表的初始框选目标框在 0 延时下就已经与目标相距较远，随着延时时长增加，该目标框与目标之间的距离也越来越大；而蓝色框代表的本文算法，虽然在延时时长增加后对目标的覆盖面积有较明显下降，但尚能捕捉目标。

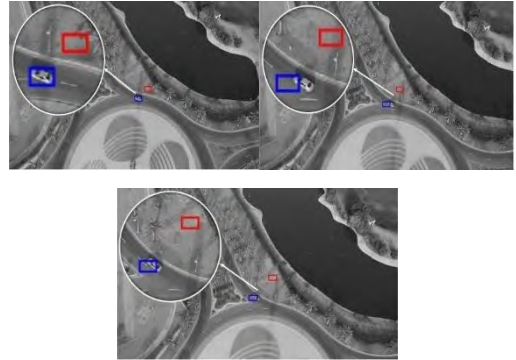


图 13 工况 3 不同延时结果对比实验图

Fig.13 Comparison experiment diagram of different delay results in condition 3

为统计算法的运行时间以验证其实时性：在各

个工况下分别运行 3 次算法，并统计其平均运行时间结果如表 3 所示。

表 3 不同工况延时误差补偿算法运行时间（单位：秒）

Table. 3 Running time of error compensation algorithm for each working conditions (s)

	工况 1	工况 2	工况 3
第一次	0.1120	0.1898	0.2136
第二次	0.1032	0.1863	0.2185
第三次	0.0985	0.1862	0.2125
平均值	0.1046	0.1874	0.2149

随着工况复杂程度的增加，算法的运行时间也随之提高，但三个工况下的最大运行时间仅有 0.2185 秒。

从以上三种工况的实验结果分析，分别对比工况 1、3 以及工况 2、3 可得，无人机盘旋状态和目标运动状态都会对算法的估计精度造成影响；而对比工况 1、2 可得，目标运动状态对算法估计精度的影响要更胜一筹。而综合对比三种工况，随着工况复杂程度的增加，算法的精度和实时性会随之下降，但精度上算法输出的目标框始终能捕捉到目标，可为后续跟踪算法提供可靠的参考坐标；实时性上虽有所降低，但算法耗时远低于手动框选操作耗时，对算法的整体运行影响甚小。

4 结论

本文通过对单台无人机捕获的图像进行目标搜寻及定位并解析坐标，将获取的目标信息共享给其他无人机，他机根据该共享信息并使用 YOLOv3 算法来完成对被跟踪目标的搜寻与定位，实现了多架无人机对目标的快速搜寻及定位，提高了任务效率；同时针对目标搜寻及定位过程中由人员操作以及通信延时所带来的延时误差问题，通过 KCF 算法、机载相机视频帧率以及通信时间差对误差大小进行估计并补偿，来获得更精确的定位结果。实验结果表明，所设计的选定方案能够提高多机的目标选定效率，并有效减小在选定过程中由操作和通信传输给目标坐标所带来的延时误差。

在下一步研究中，还需要进一步考虑共享目标信息的类型以及在复杂环境下存在多个相似物体干扰的工况中他机如何利用目标检测算法准确搜寻定位待跟踪目标。

参考文献

[1] 邸斌,周锐,董卓宁.考虑信息成功传递概率的多无人机协同目标最优观测与跟踪[J].控制与决策,2016,31(4):616-622.

[2] Oh H, Kim S. Persistent standoff tracking guidance using constrained particle filter for multiple UAVs[J]. Aerospace science and technology, 2019, 84(JAN.):257-264.

[3] 汪亮,王文,王禹又,侯松林,乔裕哲,吴天珩,陶先平.强化学习方法在通信拒止战场仿真环境中多无人机目标搜寻问题上的适用性研究[J].中国科学:信息科学,2020,50(3):375-395.

[4] 沈东,魏瑞轩,祁晓明,关旭宁.基于 MTPM 和 DPM 的多无人机协同广域目标搜索滚动时域决策[J].自动化学报,2014,40(7):1391-1403.

[5] 屈耀红,张峰,谷任能,等.基于距离测量的多无人机协同目标定位方法[J].西北工业大学学报,2019,37(2):63-69., available: <http://www.atm.com>, May 10, 1991.

[6] Redding J D, McLain T W, Beard R W, et al. Vision-based target localization from a fixed-wing miniature air vehicle[C]//American Control Conference, 2006.2006: 2862-2867.

[7] 何文涛.无人机定位技术及其应用[J].光电技术应用,2003,18(5):3-6

[8] Herekoglu O, Hasanazade M, Saldiran E, et al. Flight Testing of a Multiple UAV RF Emission and Vision Based Target Localization Method[C]//AIAA Scitech 2019 Forum. 2019: 1-18.

[9] Hou Y, Yu C. Autonomous target localization using quadrotor[C]//The 26th Chinese Control and Decision Conference (2014 CCDC). IEEE, 2014: 864-869.

[10] Lowe D G. Distinctive Image Features from Scale-Invariant Keypoints[J]. International Journal of Computer Vision, 2004, 60(2):91-110.

[11] 蔡国榕,李绍滋,吴云东,苏松志,陈水利.一种透视不变的图像匹配算法[J].自动化学报,2013,39(7):1053-1061.

[12] 杨亚飞,郑丹晨,韩敏.一种基于多尺度轮廓点空间关系特征的形状匹配方法[J].自动化学报,2015,41(8):1405-1411.

[13] Henriques J F, Caseiro R, Martins P, et al. High-speed tracking with kernelized correlation filters[J]. IEEE transactions on pattern analysis and machine intelligence, 2014, 37(3): 583-596.

[14] 刘巧元,王玉茹,张金玲,殷明浩.基于相关滤波器的视频跟踪方法研究进展[J].自动化学报,2019,45(2):265-275.

[15] 孟磊,杨旭.目标跟踪算法综述[J].自动化学报,2019,45(7):1244-1260.

[16] Danelljan M, Häger G, Khan F, et al. Accurate scale estimation for robust visual tracking[C]//British Machine Vision Conference, Nottingham, September 1-5, 2014. BMVA Press, 2014.

[17] Danelljan M, Bhat G, Shahbaz Khan F, et al. Eco: Efficient convolution operators for tracking[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 6638-6646.

[18] Bolme D S, Beveridge J R, Draper B A, et al. Visual object tracking using adaptive correlation filters[C]//2010 IEEE computer society conference on computer vision and pattern recognition. IEEE, 2010: 2544-2550.