

doi: 10.6046/gtzyyg.2020227
引用格式: 谭熊,王晶磊,孙一帆. 基于特征的无人机载视频运动目标快速检测方法[J]. 国土资源遥感,2021,33(2):27-32.
(Tan X, Wang J L, Sun Y F. A fast detection method for moving targets in UAV video based on the feature extraction [J]. Remote Sensing for Land and Resources ,2021,33(2):27-32.)

基于特征的无人机载视频运动目标快速检测方法

谭 熊, 王晶磊, 孙一帆
(战略支援部队信息工程大学, 郑州 450001)

摘要: 针对无人机载视频序列影像动态背景以及常用检测方法处理效率不高等问题,提出了一种基于特征的无人机载视频运动目标快速检测方法。该方法主要包括预处理、ORB 特征提取(oriented fast and rotated BRIEF, ORB)、渐进一致采样(progressive sample consensus, PROSAC)特征精匹配、全局运动估计与全局运动补偿、运动目标初检测和形态学后处理等 6 个步骤。通过 2 组无人机载视频数据实验研究表明:对不同传感器获取的不同帧率和不同尺寸大小的视频序列影像,采用本方法的运动目标检测结果较好,计算效率最高,且能够满足实时处理需求。
关键词: 视频序列影像; ORB 特征; PROSAC 精匹配; 运动目标检测
中图法分类号: TP 751 **文献标志码:** A **文章编号:** 1001-070X(2021)02-0027-06

0 引言

无人机载视频运动目标检测跟踪是计算机视觉领域的重要研究内容之一,已广泛应用于战场动态监测、目标侦察与打击、智能交通、灾害预警以及城市调查等军事和民用领域。无人机在获取地面目标的过程中往往存在抖动现象,同时无人机自身也在快速运动,这些因素导致了无人机视频中地面背景的不断变化,从而增加了运动目标检测跟踪的难度。无人机视频运动目标检测跟踪是一种动态场景下的运动目标检测跟踪,是计算机视觉领域的研究重点和难点。针对动态场景下的运动目标检测跟踪,众多研究学者展开了大量的研究工作,如简单的两帧差分及其改进方法(如三帧差分)^[1-2],基于区域匹配的方法^[3],基于光流技术的方法^[4-6]以及基于快速区域卷积神经网络(faster region-based convolutional neural networks, Faster R-CNN)^[7]、SSD(single shot multibox detector, SSD)^[8]、YOLO(you only look once, YOLO)^[9]等深度学习方法。在这些方法中,帧差法及其改进方法计算效率较高,但精度较低;基于区域匹配的方法往往需要目标的先验知识;基于光流技术的方法计算效率较低;基于深度学习的方法需要大量的样本来训练模型,精度较高但效率较低。在无人机视频运动目标检测跟踪方

面,国内外学者也开展了大量的研究工作,任霞^[10]提出了改进核相关滤波跟踪方法和改进LK光流进行无人机视频运动目标跟踪;针对目标变形和遮挡问题,陈挺^[11]提出了基于动态稀疏投影与目标外观模型、目标外观概率模型优化以及多层数据关联目标跟踪方法;针对目标检测跟踪实时性的问题,董晶等^[12]提出了基于特征匹配的运动目标检测法。在无人机视频运动目标检测跟踪的广泛应用中,运动目标检测的准确性和实时性是非常重要的一个因素,如在战场目标动态监测过程中,战场态势稍纵即逝,如不能及时发现、跟踪目标,很可能将失去战斗的意义。因此,对于在线实时的无人机视频运动目标检测跟踪,还有待更进一步的研究和完善。

针对无人机载视频运动目标检测准确性和实时性的问题,本文提出了一种基于特征的无人机载视频运动目标快速检测方法,采用ORB特征提取(oriented fast and rotated BRIEF, ORB)和渐进一致采样(progressive sample consensus, PROSAC)精匹配算法进行全局运动估计和运动补偿,然后采用帧差法和自适应形态学操作进行运动目标检测,检测精度和效率较高,能够满足实时处理需求。

1 研究方法

本文所提方法的基本思路是:首先,对获取的

收稿日期: 2020-07-24; 修订日期: 2021-10-27
基金项目: 高分重大专项项目(编号: GFZX04040202-04)资助。
第一作者: 谭 熊(1986-),男,讲师,主要研究方向为高光谱遥感、计算机视觉、目标工程等。Email: kjadetxt@163.com。

视频序列影像进行直方图均衡化等亮度归一化预处理,并利用 ORB 算法进行特征提取,在常用匹配方法的基础上采用 PROSAC 精匹配剔除错误匹配;然后,采用六参数模型根据最小二乘原理进行全局运动估计与运动补偿,将补偿后的序列影像采用帧差法进行运动目标初检测;最后,通过形态学操作等后处理实现最终的运动目标检测。本文所提方法的处理流程如图 1 所示。

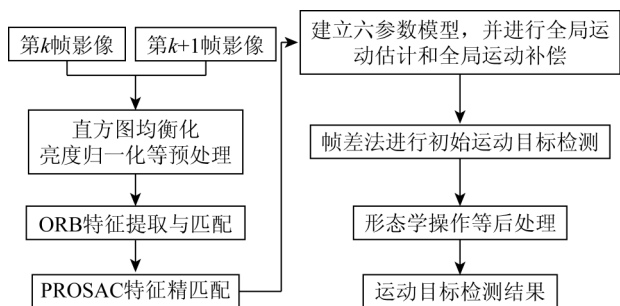


图 1 本文方法的处理流程

Fig. 1 Processing flowchart of processed method

1.1 ORB 特征提取与匹配

ORB^[13] 特征检测算法采用改进的 FAST 关键特征点检测和改进的 BRIEF 特征描述子进行关键特征检测,具有尺度和旋转不变性,且检测效率优于传统的 Harris, SIFT (scale - invariant feature transform, SIFT) 以及 SURF (speeded up robust features, SURF) 等特征检测方法,可应用于实时特征检测场景中。

1.1.1 FAST 关键特征点检测

FAST 关键特征点检测^[14] 的基本思路是以影像中某一像素点 (候选特征点) 为基准,计算该候选特征点与其周围像素点 (一般为 16 个点) 的图像灰度差值,如果周围邻域内有足够多的像点满足下式的阈值要求,则认为该候选特征点为一个特征点,即

$$N = \sum_{x \in \text{circle}(p)} |I(x) - I(p)| > \varepsilon_d, \quad (1)$$

式中: $I(x)$ 为周围邻域圆周上任意一点的灰度值; $I(p)$ 为候选特征点的灰度值; ε_d 为灰度阈值。如果 $N > \varepsilon_d$, 且数量大于周围像素点的 3/4 时,则 p 为关键特征点。FAST 关键特征点检测示意图如图 2 所示。在 ORB 方法中,为了提高特征提取效率,通常采用周围 9 个点进行计算,即 FAST-9。在计算出关键特征点后,采用非极大值抑制法消除局部较密集特征点。另外,FAST 关键特征点检测本身不具有尺度不变和旋转不变性。在 ORB 方法中,采用图像金字塔形式实现特征点的尺度不变性;采用灰度质心法来计算关键特征点的特征方向以实现特征点的旋转不变性。灰度质心法的基本思路是:首先在邻域范围内计算每个像素以灰度值为权重的灰度质

心,然后连接关键特征点与灰度质心,所构成的方向即为该关键特征点的特征方向。

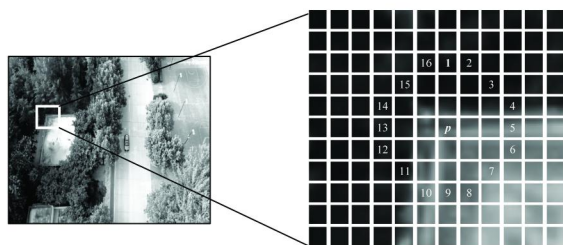


图 2 FAST 关键特征点检测示意图

Fig. 2 Schematic diagram of FAST key feature point detection

1.1.2 BRIEF 特征描述符

BRIEF 特征描述符^[15] 是一种二进制描述符,其核心思想是从关键点 p 周围随机选取 N 个点 (一般为 128 个点),然后计算每个点对中前一个点与后一个点的灰度差值,如果大于 0,则取值为 1,如果小于等于 0,则取值为 0。其公式为:

$$T(P(p_1, p_2)) = \begin{cases} 1 & I_{p_1} - I_{p_2} > 0 \\ 0 & I_{p_1} - I_{p_2} \leq 0 \end{cases}, \quad (2)$$

式中: $P(p_1, p_2)$ 表示一个由 (p_1, p_2) 2 个关键特征点组成的点对; I_{p_1} 为 p_1 点的灰度值。

当计算完所有点对的 T 值后,生成一个 128 维的二进制编码,该编码即为该特征点的特征描述符。在得到特征描述符后,可用简单的汉明距离实现特征匹配。然而,常用的 BRIEF 特征描述符在进行特征匹配时不具备旋转不变性,造成了大量的误匹配,针对该问题,ORB 方法采用具有旋转不变性的改进方法——rBRIEF 特征描述符。

1.1.3 特征匹配

特征匹配的方法很多,最简单的是距离匹配法,包括汉明距离和暴力匹配器。

汉明距离匹配法是计算 2 个特征点特征描述符 (二进制编码字符串) 对应位置不同的字符个数,若相差个数越少,则距离越近,即为特征匹配点。暴力匹配器是计算某一个特征点描述符与其他所有特征点描述符之间的距离,取距离最小的特征点作为匹配点。汉明距离法和暴力匹配器方法简单,计算效率高,但是往往存在大量的错误匹配,因此需要采用一定的方法消除错误匹配。

1.2 PROSAC 特征精匹配

在特征精匹配方法中,随机抽样一致算法 (random sample consensus, RANSAC) 算法^[16] 是一种常用的鲁棒方法,但由于其高度随机采样数据的特点导致了计算量较大,处理精度和效率较低。针对该问

题,本文采用 RANSAC 改进方法——PROSAC 法^[17]来剔除误匹配,实现特征精匹配。PROSAC 算法首先在初始匹配结果中计算最小距离和次小距离的比值,通过质量因子衡量匹配的质量;然后对数据质量按照由高到底的顺序进行排序;最后根据高质量的匹配子集数据进行单应矩阵模型参数估计。质量因子的计算公式为:

$$\gamma = \frac{1}{\beta d_{\min}}, \quad (3)$$

$$\beta = \frac{d_{\min}}{d_{\min 2}}, \quad (4)$$

式中 β 为最小距离 $d_{\min}$ 与次小距离 $d_{\min 2}$ 的比值。

1.3 全局运动估计与运动补偿

无人机搭载的视频传感器的运动引起了视频序列影像的全局背景位移,视频传感器的位置和姿态的变化可以分解为传感器的平移和旋转。相应地在二维图像平面上,图像上的任一像素点 $p(x, y)$ 经过时间 Δt 后,运动到另一点 $p'(x', y')$,该运动可以用旋转加平移来表示为:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a_1 & a_2 \\ a_4 & a_5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} a_3 \\ a_6 \end{pmatrix}, \quad (5)$$

式中: $\begin{pmatrix} a_1 & a_2 \\ a_4 & a_5 \end{pmatrix} = M$, 为旋转分量; $\begin{pmatrix} a_3 \\ a_6 \end{pmatrix} = b$, 为平移分量。表示成线性方程组为:

$$\begin{cases} x' = a_1 x + a_2 y + a_3 \\ y' = a_4 x + a_5 y + a_6 \end{cases}. \quad (6)$$

式(5)称为运动估计的六参数模型。在完成特征提取和匹配后,利用至少 3 个特征点对根据最小二乘原理进行模型参数解算,当解算出参数模型中的 6 个参数 ($a_1, a_2, a_3, a_4, a_5, a_6$) 后,就可以根据六参数模型对当前帧进行重采样,进而得到补偿后的图像,实现全局运动估计和全局运动补偿。

1.4 运动目标的检测

经过全局运动补偿后,消除了无人机视频序列影像背景的位移,直接采用帧差法进行运动目标的检测。设 $I_n(x, y)$ 和 $I_{n+1}(x, y)$ 为连续 2 帧影像在 (x, y) 处的灰度值,帧差法的计算公式为:

$$P_n(x, y) = \begin{cases} 1 & |D_n(x, y)| > T \\ 0 & |D_n(x, y)| \leq T \end{cases}, \quad (7)$$

$$D_n(x, y) = I_{n+1}(x, y) - I_n(x, y), \quad (8)$$

式中: $P_n(x, y)$ 为二值化图像; $D_n(x, y)$ 为灰度差值; T 为阈值。

在实际情况下,由于受光照、风以及噪声等外界

因素的干扰,帧差法所得结果并不能完全消除序列影像的背景位移,往往存在背景噪声等影响。为了消除该影响,本文采用形态学滤波操作,对运动目标初始检测结果进行后处理,进而得到最终的运动目标。

2 实验与分析

采用 2 组真实无人机载视频数据进行实验,实验内容包括运动目标检测结果和计算效率 2 部分内容,计算效率实验未采用图形处理器 (graphics processing unit, GPU) 进行加速计算。实验硬件环境为 16 G 内存, i9 处理器, 64 位操作系统, 软件环境为 Python3.6 和 OpenCV3.4.2。

2.1 实验数据

实验采用 2 组数据: 数据一为大疆御 2 商用无人机获取的地面车辆运动目标视频, 视频图像大小为 3 840 像素 $\times$ 2 160 像素, 帧率为 30 帧/s, 视场中树木较多, 容易受光照和噪声等影响。采用连续 2 帧影像 (图 3) 进行运动估计和运动目标检测实验。



(a) 第 k 帧影像

(b) 第 $k+1$ 帧影像

图 3 大疆御 2 无人机视频序列影像

Fig. 3 *Dji Royal 2 UAV video sequence imagery*

数据二为某型无人飞艇获取的地面车辆目标视频序列影像, 图像大小为 305 像素 $\times$ 245 像素, 帧率为 25 帧/s。由于飞行高度较高, 视场范围较大, 运动目标较小, 给运动目标检测带来一定的困难。其连续 2 帧原始视频序列影像如图 4 所示。



(a) 第 k 帧影像

(b) 第 $k+1$ 帧影像

图 4 某型无人飞艇视频序列影像

Fig. 4 *Unmanned airship video sequence imagery*

2.2 运动目标检测实验结果与分析

按照图 1 所示的处理流程分别对 2 组视频序列影像进行直方图均衡化预处理、特征提取与匹配、运动估计与运动补偿、帧差法运动目标检测。直方图均衡化预处理的目的是消除光照、噪声等

变化引起的同一目标灰度值变化较大的影响。在实验参数设置方面,主要有帧差法二值化阈值、形态学形状及其核大小 2 个参数: ①帧差法二值化阈值 T 的选取: 设阈值范围为 10 ~ 200, 每间隔 5 进行一次帧差法二值化实验, 得到最佳的二值化分割阈值为 50; ②形态学模板及其核大小的选择: 采用 3×3 , 5×5 和 7×7 等不同核大小的矩形模板进行形态学开运算, 选取效果最好的 5×5 矩形模板进行形态学处理, 得到数据一的实验结果如图 5—图 9 所示。

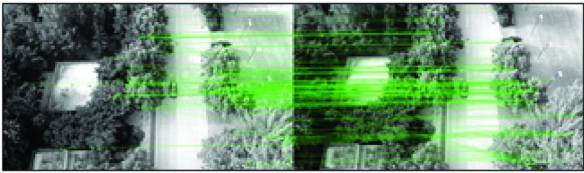


图 5 数据一粗匹配结果
Fig. 5 Result of coarse matching of data 1

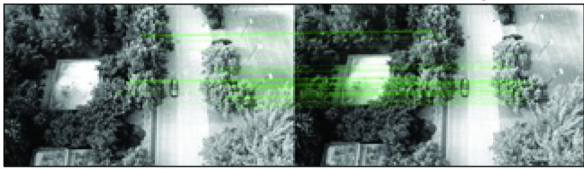
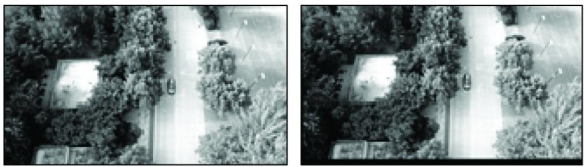
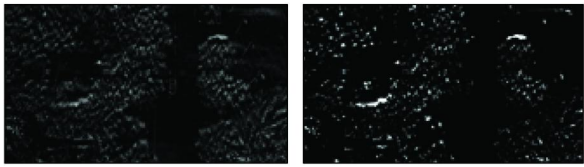


图 6 数据一精匹配结果
Fig. 6 Result of fine matching of data 1



(a) 第 k 帧影像 (b) 运动补偿后影像
图 7 数据一运动补偿结果
Fig. 7 Result of motion compensation of data 1



(a) 直接差分法初始检测结果 (b) 形态学处理后运动目标检测结果
图 8 数据一原始影像帧差法运动目标检测结果
Fig. 8 Results of moving target detection by original image frame using difference method of data 1



(a) 初始检测结果 (b) 形态学处理后运动目标检测结果
图 9 数据一本文所提方法运动目标检测结果
Fig. 9 Result of moving target detection results of the proposed method of data 1

数据二实验参数设置的方法与数据一相类似, 具体参数为: 帧差法二值化阈值 T 设为 30, 采用核大小为 3×3 的矩形模板进行形态学腐蚀运算后处理, 实验结果如图 10—图 14 所示。

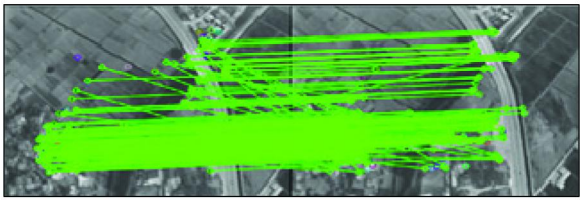


图 10 数据二粗匹配结果
Fig. 10 Result of coarse matching of data 2

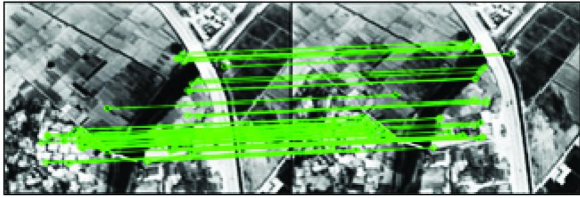
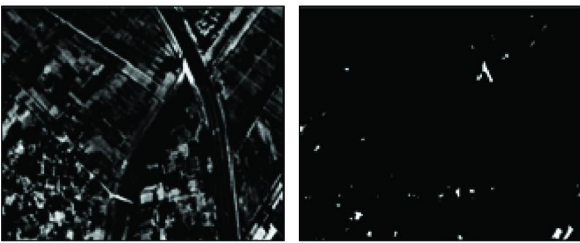


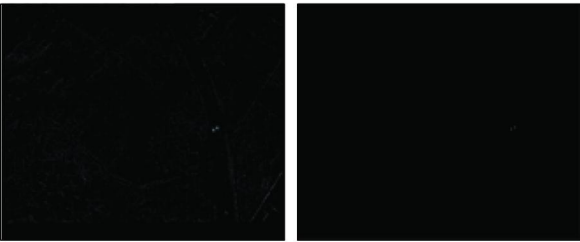
图 11 数据二精匹配结果
Fig. 11 Result of fine matching of data 2



(a) 第 k 帧影像 (b) 运动补偿后影像
图 12 数据二运动补偿后结果
Fig. 12 Result of motion compensation of data 2



(a) 直接差分法初始检测结果 (b) 形态学处理后运动目标检测结果
图 13 数据二原始影像帧差法运动目标检测结果
Fig. 13 Results of moving target detection by original image frame using difference method of data 2



(a) 初始检测结果 (b) 形态学处理后运动目标检测结果
图 14 本文所提方法运动目标检测结果
Fig. 14 Result of moving target detection results of the proposed method

分析以上 2 组实验结果,可以得出如下结论:

1)图 5—图 6 以及图 10—图 11 的结果表明,采用传统的汉明距离进行特征匹配,存在较多的错误匹配;而采用 PROSAC 精匹配方法能够全部消除错误匹配,得到较高精度的匹配结果。

2)从图 8 和图 13 的实验结果可以看出,采用直接差分法并不能检测出动态场景下的运动目标。

3)从图 9 和图 14 的实验结果可以看出,不管视频场景中运动目标大小如何,采用本文所提方法均能较好地检测出运动目标,通用性较好,在此基础上可进一步实现运动目标的跟踪。

2.3 运动目标检测效率实验与分析

由于 SIFT 和 SURF 算法特征检测精度较高而得到广泛应用,因此本节主要将基于 ORB 特征提取的运动目标检测与基于 SIFT 和 SURF 特征提取的运动目标检测计算效率进行对比分析,2 组数据的实验结果如表 1 所示,表中数据为运行 10 次的平均值。

表 1 运动目标检测效率

Tab. 1 Moving target detection efficiency (“)

方法	时间	
	数据一	数据二
SIFT + RANSAC	132.530	0.078
SIFT + PROSAC	134.030	0.093
SURF + RANSAC	23.750	0.047
SURF + PROSAC	23.690	0.047
ORB + RANSAC	0.344	0.016
ORB + PROSAC	0.344	0.016

分析表 1 的实验结果,可以得出如下结论:

1)运动目标检测的效率与视频序列影像大小有一定关系,通常情况下:影像尺寸越大,计算时间越长,效率越低;影像尺寸越小,计算时间越短,效率越高。

2)基于 SIFT 和 SURF 特征提取的运动目标检测计算时间明显高于基于 ORB 特征提取运动目标检测方法。

3)对于帧率为 30 帧/s 的数据一,ORB 基本能够实现实时处理需求;对于帧率为 25 帧/s 的数据二,ORB 能够较好满足实时处理需求。

综合上述实验结果,可以得出:本文所提基于 ORB 特征的无人机载视频运动目标检测方法的运动目标检测效果较好,计算效率较高,通用性较强,能够满足实时处理需求。

3 结论

本文针对动态场景下无人机载视频序列影像检

测精度不高、计算效率较低的问题,提出了基于 ORB 特征的无人机载视频运动目标检测方法。该方法首先采用直方图均衡化进行亮度归一化预处理,在此基础上采用 ORB 特征提取和 PROSAC 精匹配方法进行前后两帧影像匹配,再采用六参数模型进行视频序列影像背景的全局运动估计和运动补偿,最后采用帧差法和形态学操作进行运动目标检测。通过 2 组不同无人机载视频传感器获取的地面运动车辆数据进行实验,实验结果表明:

- 1)利用 ORB 特征提取后再进行 PROSAC 精匹配的方法,可实现相邻两帧无人机视频序列影像高精度配准。
- 2)本文所提方法能够实现动态场景下不同尺寸运动目标的检测,检测效果较好,算法通用性较强。
- 3)本文所提方法计算效率优于基于 SIFT 和 SURF 特征提取的运动目标检测方法,且计算时间基本满足无人机视频实时处理需求。

同时在本文研究过程中还存在一定不足之处:一是未采用 GPU 并行计算来处理更大尺寸的无人机视频数据;二是没有考虑运动目标的遮挡。这些问题是无人机视频运动目标检测跟踪的重点和难点内容,有待进一步研究。

参考文献(References):

[1] Mu Y W,Guo C H,Xin Y D. A new interframe difference algorithm for moving target detection[C]//IEEE International Congress on Image and Signal Processing,2010(1):285-289.

[2] 熊 英. 基于背景和帧间差分法的运动目标提取[J]. 计算机时代,2014(3):38-41.

Xiong Y. Moving object extraction based on background difference and frame difference method[J]. Computer Era,2014(3):38-41.

[3] Kengo M,Takashi S,Shoji Y, et al. Moving-object detection method for moving cameras by merging background subtraction and optical flow methods[C]//IEEE Global Conference on Signal and Information Processing,2017:383-387.

[4] 谭 熊. 基于光流分析的无人机视频运动目标检测与跟踪[D]. 郑州:解放军信息工程大学,2011.

Tan X. UAV video moving target detection and tracking based on optical flow analysis[D]. Zhengzhou:Information Engineering University,2011.

[5] 王朋月. 基于全局运动估计补偿的目标检测算法研究[D]. 北京:北京工业大学,2018.

Wang P Y. Research on target detection algorithm based on global motion estimation and compensation[D]. Beijing:Beijing Industry University,2018.

[6] 卢裕秋,孙金玉,马世伟. 基于深度卷积神经网络的运动目标检测方法[J]. 系统仿真学报,2019,31(11):2275-2280.

Lu Y Q,Sun J Y,Ma S W. Moving object detection based on deep

convolutional neural network [J]. Journal of System Simulation , 2019 , 31 (11) : 2275 – 2280 .

[7] Ren S , He K , Gir S R , et al . Faster R – CNN : Towards real – time object detection with region proposal networks [J]. IEEE Trans Pattern Anal Mach Intell , 2015 , 39 (6) : 1137 – 1149 .

[8] Redmon J , Divvala S , Girshick R , et al . You only look once : Unified , real – time object detection [C] // IEEE Conference on Computer Vision and Pattern Recognition . IEEE , 2016 : 779 – 778 .

[9] Liu W , Anguelov D , Erhan D , et al . SSD : Single shot multi – box detector [C] // European Conference on Computer Vision , Amsterdam . IEEE , 2016 : 21 – 37 .

[10] 任 霞 . 基于无人机视频的运动目标检测与跟踪方法研究 [D]. 成都 : 电子科技大学 , 2020 .

Ren X . Research of detection and tracking based on the moving targets in UAV video [D]. Chengdu : University of Electronic Science and Technology of China , 2020 .

[11] 陈 挺 . 无人机对地运动目标跟踪技术研究 [D]. 西安 : 西北工业大学 , 2018 .

Chen T . Moving objects tracking in aerial videos [D]. Xi ' an : Northwestern Polytechnical University , 2018 .

[12] 董 晶 , 傅 丹 , 杨 夏 . 无人机视频运动目标实时检测及跟踪 [J]. 应用光学 , 2013 , 34 (2) : 255 – 259 .

Dong J , Fu D , Yang X , et al , Real – time moving object detection and tracking by using UAV videos [J] , Applied Optics , 2013 , 34 (2) : 255 – 259 .

[13] Rublee E , Rabaud V , Konolige K , et al . ORB : An efficient alternative to SIFT or SURF [C] // Proceedings of IEEE International Conference on Computer Vision , Washington , USA , 2011 : 2564 – 2571 .

[14] Mair E , Hager G D , Burschka D , et al . Adaptive and generic corner detection based on the accelerated segment test [C] // Proceeding of European Conference on Computer Vision , Crete , Greece , 2010 , 63 (12) : 183 – 196 .

[15] Calonder M , Lepetit V , Fua P . BRIEF : Binary robust independent elementary features [C] // Proceeding of European Conference on Computer Vision , Crete , Greece , 2010 , 63 (14) : 778 – 792 .

[16] Fischler M A . Random sample consensus : A paradigm for model fitting with applications to image analysis and automated cartography [J]. Readings in Computer Vision , 1987 : 726 – 740 .

[17] Chum O . Matching with PROSAC — Progressive Sample Consensus [J]. Processing of Conference on CVPR , 2005 , 1 (2) : 220 – 226 .

A fast detection method for moving targets in UAV video based on the feature extraction

TAN Xiong , WANG Jinglei , SUN Yifan

(Strategic Support Force Information Engineering University , Zhengzhou 450001 , China)

Abstract: Aiming at tackling the problem of the low efficiency of common detection methods and the dynamic background of UAV – based video sequences imagery , this paper proposes a fast detection method for moving targets in UAV video based on the feature extraction . The method mainly includes six steps , i . e . , preprocessing , ORB feature extraction , PROSAC feature fine matching , global motion estimation and global motion compensation , initial detection of moving target and morphological post – processing . The results of two video experiments carried by UAVs show that the moving target detection results of the proposed method in this paper are better , the computational efficiency is the highest , and hence this method can meet the requirements of real – time processing .

Keywords: video sequential imagery ; ORB feature ; PROSAC fine matching ; moving target detection

(责任编辑 : 李 瑜)