

引用格式:张艳国,李擎,于飞,等.基于多特征融合的空对地目标检测和追踪方法[J].电光与控制,2019,26(6):6-11. ZHANG Y G, LI Q, YU F, et al. An air-to-ground target detecting and tracking method based on multi-feature fusion[J]. Electronics Optics & Control, 2019, 26(6):6-11.

## 基于多特征融合的空对地目标检测和追踪方法

张艳国, 李 擎, 于 飞, 刘恒志

(北京信息科技大学高动态导航技术北京市重点实验室,北京 100101)

**摘 要:** 随着无人机的广泛应用,无人机对地面目标检测和追踪技术成为研究的热点。针对光流法对运动目标检测和追踪出现误检、丢失的问题,应用运动目标纵向的边缘特征配合光流法的方法完成目标检测,应用 SIFT 与光流法实现目标追踪,并提出恒速模型与光流块匹配的方法解决追踪过程中建筑物等对目标的遮挡问题。实验结果表明,所提方法能够对目标进行有效的识别与追踪,并具有较强的鲁棒性。

**关键词:** 无人机; 空对地目标检测; 光流法; SIFT; 目标追踪

**中图分类号:** TP391.4 **文献标志码:** A **doi:**10.3969/j.issn.1671-637X.2019.06.002

## An Air-to-Ground Target Detecting and Tracking Method Based on Multi-feature Fusion

ZHANG Yan-guo, LI Qing, YU Fei, LIU Heng-zhi

(Beijing Key Laboratory of High Dynamic Navigation Technology, Beijing Information Science & Technology University, Beijing 100101, China)

**Abstract:** With the wide application of Unmanned Aerial Vehicles (UAVs), using UAVs for detecting and tracking the ground targets has become a research hotspot. Aiming at the problem of false detection and loss of target during the detection and tracking of moving objects by the optical flow method, the target detection is performed by using the longitudinal edge features of the moving object in conjunction with the optical flow method. Simultaneously, Scale-Invariant Feature Transform (SIFT) and optical flow method are used to realize target tracking. A method of matching the constant-speed model with the optical flow block is proposed to solve the problem of occlusion of the target by buildings during tracking. The experimental results show that the method can implement effective target identification and tracking, and has strong robustness.

**Key words:** UAV; air-to-ground target detection; optical flow method; SIFT; target tracking

### 0 引言

无人机又称为无人驾驶飞行器,是一种具有可遥控、自动、半自主、全自主飞行能力的飞行器。常见的无人机主要分为固定翼无人机、无人直升机和多旋翼无人机。随着控制技术、传感器技术和计算机视觉领域的快速发展,多旋翼无人机成为航拍领域的新星,尤其是四旋翼无人机和八旋翼无人机受到了航拍领域的青睐<sup>[1]</sup>。

目标检测和追踪技术是无人机航拍领域的重要研

究方向。目标检测与追踪技术作为计算机视觉领域的关键技术,受到了各界学者的广泛关注<sup>[2]</sup>。在无人机航拍领域中,为了实现追踪拍摄,目标检测和追踪必不可少,因此,目标检测和追踪技术是无人机航拍领域的重要研究方向<sup>[3-4]</sup>。另外,由于航拍视频画面较大,目标在场景中所占面积较小,背景复杂,目标易受到尺度、旋转、光照和遮挡干扰以及相机抖动等影响,航拍视频中目标的检测和追踪变得尤为复杂,检测和追踪工作变得更加困难。本文针对无人机对路面车辆进行检测与追踪过程中目标物错检和被遮挡的情况展开研究。

### 1 实验条件

拍摄环境为一双向四车道主干道路,如图1所示,考虑天气和光照的情况,在早上6点半到7点进行视频获取。无人机飞行高度分别为100 m,120 m以及

收稿日期:2018-07-24

修回日期:2018-08-24

基金项目:国家自然科学基金(61471046);北京市自然科学基金(4172022);北京市科技计划课题(Z161100003016011)

作者简介:张艳国(1991—),男,河北石家庄人,硕士生,研究方向为导航、制导与控制技术。

150 m, 视野中道路长度分别为 190 m, 228 m 和 285 m。无人机主要技术参数见表 1。

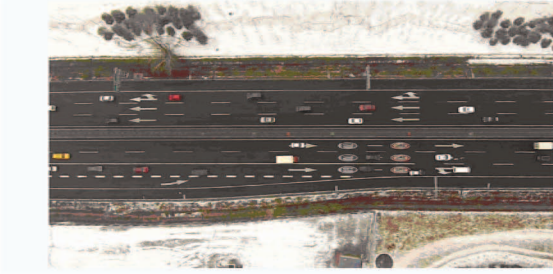


图1 实验道路图

Fig.1 Experimental road

表1 无人机(大疆 Phantom 2)主要技术参数

Table 1 Main technical parameters of UAV (DJI Phantom 2)

参数	指标
悬停精度/m	垂直 0.8, 水平 2.5
最大飞行速度/(m · s <sup>-1</sup> )	15(不推荐)
最大飞行时间/min	25
工作温度/°C	-10 ~ 50
最大旋转角速度/(° · s <sup>-1</sup> )	200
最大俯仰角度/(°)	35
最大上升/下降速度/(m · s <sup>-1</sup> )	6/2

## 2 实验方法

### 2.1 目标检测

昆虫视觉导航主要是依赖物体在视网膜上的运动产生的光流,它们可以获取丰富的空间特征信息。按照此原理,在图像处理中像素的移动称为光流法,经常用于目标检测与追踪,可分为稠密光流法和稀疏光流法<sup>[5]</sup>。本节将使用稀疏光流法进行目标检测,并融合边缘检测算法,解决检测过程中由于亮度不恒定和像素特征较少引起的误检。

#### 2.1.1 LK 光流法

Lucas-Kanade (LK) 光流法是常用的一种稀疏光流法,它假定在所考虑像素的局部邻域中流动基本上是恒定的,并且通过使用最小平方准则来求解该邻域中所有像素的基本光流方程。像素的光流方程可以假定为对于以  $p$  为中心的区域  $\Omega$  内的所有像素,也就是说,局部图像光流向量  $(u, v)$  可以通过计算加权项的最小值来获得,即

$$E = \sum_{x,y \in \Omega} (I_x V_x + I_y V_y + I_t)^2 \quad (1)$$

最小化  $E$ , 可得  $A^T A V = A^T b$ , 其中,  $A = [\nabla I(X_1), \dots, \nabla I(X_n)]$ ,  $b = -[I_t(X_1), \dots, I_t(X_n)]^T$  为常数项。

上式解为

$$V = [A^T A]^{-1} A^T b \quad (2)$$

$$\begin{bmatrix} V_x \\ V_y \end{bmatrix} = \begin{bmatrix} \sum_i I_x(X_i)^2 & \sum_i I_x(X_i) I_y(X_i) \\ \sum_i I_y(X_i) I_x(X_i) & \sum_i I_y(X_i)^2 \end{bmatrix}^{-1} \cdot \begin{bmatrix} \sum_i I_x(X_i) I_t(X_i) \\ \sum_i I_y(X_i) I_t(X_i) \end{bmatrix} \quad (3)$$

每辆车的速度包括纵向和横向两个方向。在无人机视频中,大多数横向运动属于相机运动。为了使误差最小化,只考虑沿着道路纵向的车辆速度,可以表示为

$$V_o = V_x \cos \vartheta + V_y \sin \vartheta \quad (4)$$

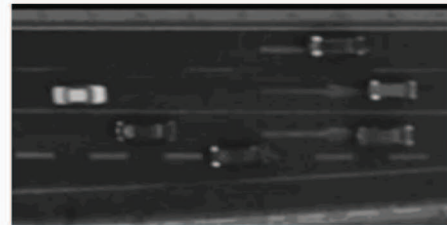
式中,  $\vartheta$  表示图像方向,即  $x$  和  $y$  方向的光流速度与道路纵向平道线的夹角。

虽然可以得到方程中的每个像素的速度,但由于像素级的光流不稳定,很难从无人机的图像序列的光流结果中精确地直接获得每辆车的速度。因此, LK 光流法主要用来检测每辆车的形状,并且  $V_o$  的值将根据式(5)中适当的阈值转换成二进制值。

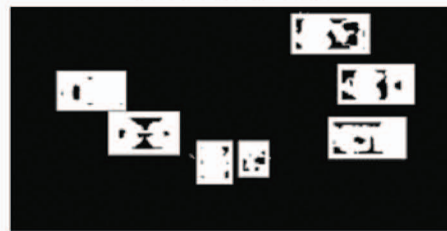
$$F(i, j) = \begin{cases} 1 & \beta \geq V_o(i, j) \geq \gamma \\ 0 & \text{其他} \end{cases} \quad (5)$$

式中,  $\beta$  和  $\gamma$  分别是该像素点光流的最大和最小速度,它们与无人机的高度和图像配准后的稳定性有关。较高的海拔导致较小的  $\gamma$ , 而较低稳定性则需要较大的  $\beta$  消除相机运动。

图2是一帧的最终转换结果。



a 150 m 高度的图像



b 形状检测结果

图2 LK 光流法对运动目标进行形状检测

Fig.2 Shape detection of moving targets with LK optical flow method

#### 2.1.2 LK 光流法融合边缘检测

根据 2.1.1 节获得的光流图像特征结果可以检测粗糙的车辆形状信息,然而,光流的结果在车辆形状检测过程中存在一些潜在的错误。首先,车辆内部的相似亮度意味着  $V = [A^T A]^{-1} A^T b$  中的矩阵  $A$  是奇异的。当车辆区域内的像素强度相似时,矩阵  $A$  将接近

于零。在这一点上的速度,式(2)看作零。因此,当较长车辆的特征较少时,光流法会将长车检测为两个分开的部分,这是一个错误。另一个错误是当两车靠近时,光流法会将其视为一辆车,如图 3 所示。

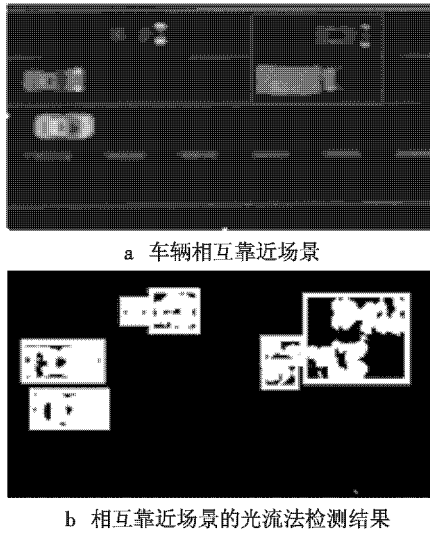


图 3 LK 光流法检测结果

Fig. 3 Detection results of LK optical flow method

针对图像特定目标的识别,有时在较为复杂的外部环境中,单纯依靠某一种特征无法实现有效的目标图像识别,由此,研究学者们提出将图像实时识别技术中的多种特征综合起来,并且通过发挥各类特征识别提取的优势,结合实际运用环境实现对目标图像识别与实时追踪效果、准确度的最优化。

目标图像识别中描述物体最为基本的特征之一是形状特征,它表现了对拍摄到的目标物体在空间结构上的一系列特性。形状特征与颜色特征的最大区别在于其不受光照影响,在光照发生变化的情况下依然能够保持原有的一致性。

针对上述问题,边缘特征被用作光流法的补充。边缘检测目的是检测数字图像中图像亮度是否急剧变化或是不是连续性的点。通过研究发现,在大多数情况下 Canny 算法对目标物边缘进行检测后所取得的结果要好于其他算法。但对应到本文中的目标图像,Canny 算法对于道路上的其他目标干扰较为敏感,比其他的算法实验结果可以得出,Prewitt 算法对图像预处理具有较好的效果。在该方法中,采用 Prewitt 算子对光流检测出的矩形区域进行边缘检测。Prewitt 算子采用  $3 \times 3$  内核与灰度图像进行邻域卷积,可以检测水平边缘和竖直边缘,用算式表示为

$$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} * A \quad (6)$$

$$G_y = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} * A \quad (7)$$

式中: \* 表示二维卷积运算;  $A$  是灰度图像,由原来的无人机视频转换而来。

$$G_o = G_x \cos \alpha + G_y \sin \alpha \quad (8)$$

只分析平行于道路的边缘方向,然后基于每个像素的  $G_o(i,j)$ ,可以生成具有合适阈值  $\delta$  的边缘二值图像,其中,  $i$  和  $j$  分别是  $x$  轴和  $y$  轴处的像素索引。 $\delta$  的值受无人机相机实验环境和操作参数的影响。二进制

$$F_e(i,j) = \begin{cases} 1 & G_o(i,j) \geq \delta \\ 0 & G_o(i,j) < \delta \end{cases}$$

边缘检测结果如图 4 所示。

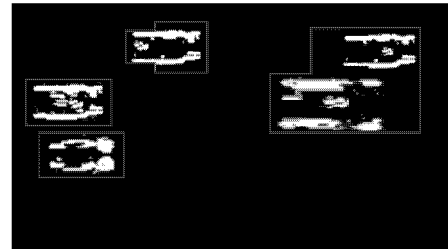


图 4 边缘检测

Fig. 4 Edge detection

### 2.1.3 检测结果

将每一帧图像进行灰度处理,采用 LK 稀疏光流法检测出车辆,Prewitt 算子再对光流法检测出的矩形区域进行边缘检测,检测出车辆平行于道路的边缘,利用检测出的边缘与光流法检测的车辆融合,就可以准确检测出运动物体,如图 5 所示。

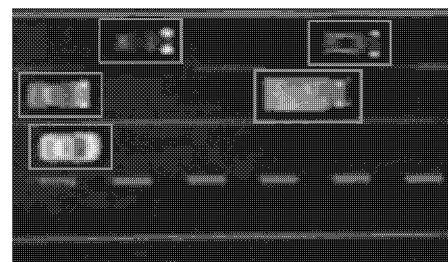


图 5 车辆检测实验

Fig. 5 Vehicle detection test

实验结果证明了所设计的特征模型对于提高无人机航拍实时图像识别的准确性效果显著。

## 2.2 目标追踪

图像目标识别与追踪包括目标实时识别与目标实时追踪两方面。首先对图像传感器所获取的数据信息予以处理,针对图像内的目标物体予以有效检测和识别,根据目标物体在当前的图像帧内的相对空间位置等数据信息进行计算。

所谓运动目标追踪即在运动目标检测的基础上,

提取要追踪物体的有效特征,利用恰当的匹配算法,在每一帧图像中寻找选定的运动目标物在图像中的位置。此处采用融合了尺度不变特征变换(SIFT)的光流法追踪无人机视频的指定运动目标。

### 2.2.1 SIFT 特征

尺度不变特征变换(SIFT)算法<sup>[6]</sup>由不列颠哥伦比亚大学的LOWE教授于1999年提出,完善于2004年,实现两幅图像的匹配。目前已被广泛应用于许多领域,如视频中物体检测、追踪,图像合成拼接,3D建模等。SIFT算法作为局部图像特征描述领域的开创性成果,算法同时对检测的目标图像进行的旋转、尺度变化、光照变化和一定视角等操作,其局部的图像特征描述算法都拥有良好的不变性,同时图像也具备极强的可区分性,较好地兼顾了局部特征描述子的两项核心指标,在图像识别追踪等领域获得了广泛运用<sup>[7]</sup>。

无人机在拍摄视频的过程中,视频中的每台车辆将随着前进而改变位置,因此,车辆角度会相应地改变,摄像头在录制视频时设置的亮度和白平衡可能会稍微改变,此外,无人机自身的运动和震动也会对拍摄的视频造成影响。由于SIFT局部特征点对于图像缩放、平移和旋转是不变的,并且对于照明变化和仿射或3D投影也是不变的,所以选用该方法用于特征点检测。尺度不变特征变换的过程为先计算图像中的局部极值,将这些极值点称为关键点。为了完成这一步,将图像中每一个像素点与不同方差的高斯函数卷积,得到不同的尺度,对于同一尺度找到局部极值,对于每一个像素求出尺度空间中的局部极值。为了提高效率,采用DOG的金字塔逼近法加快计算,即

$$\begin{cases} \sigma \nabla^2 = \frac{\partial G}{\partial \sigma} \approx \frac{G(x, y, k\sigma) - G(x, y, \sigma)}{k\sigma - \sigma} \\ G(x, y, k\sigma) - G(x, y, \sigma) \approx (k-1)\sigma^2 \nabla^2 G \\ D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) = \\ L(x, y, k\sigma) - L(x, y, \sigma) \end{cases} \quad (9)$$

DOG多尺度图像中沿着边缘方向主曲率的值会很小,垂直边缘方向主曲率的值会很大,因此可以通过计算尺度图像的Hessian值,并给定阈值筛选关键点。接下来用梯度方向直方图描述该关键点,即以关键点为中心,在其周围邻域内计算对应辐角(方向)、幅度(大小),划分区间进行统计,以值最大的区间方向作为该关键点的主方向,形成128维的空间向量,将这个向量归一化处理,可以满足光照不变性。最后,通过最小欧氏距离找到两幅图像中的对应关键点进行匹配,匹配高于阈值表示匹配完成,否则选用次最小欧氏距离进行匹配,直到完成。

具体步骤如下:

- 1) 建立尺度空间,主要是建立 Gaussian 金字塔和 DOG 金字塔;
- 2) 在尺度空间中寻找极值点,并对其进行精确定位;
- 3) 求解关键点的梯度幅值和幅角;
- 4) 对关键点进行描述,形成128维的空间向量;
- 5) 找到待匹配特征的最近邻和次近邻,从而确定匹配的候选特征,依次进行,完成匹配;
- 6) 判断匹配对是否大于阈值,若大于进行7),否则返回执行5);
- 7) 用RANSAC算法对已经匹配的特征点进行剔除,去掉错误的点对,用剩余点对得到一个新的仿射变换矩阵  $R$ ;
- 8) 用仿射变化矩阵  $R$  将上一个目标矩形框的4个顶点进行变换,得到新的目标矩形框,实现新的目标区域的定位;
- 9) 依次进行下去,直至视频流的最后一帧图像。

### 2.2.2 LK 光流法用于目标追踪

众所周知,SIFT算法是特征点研究的有力工具。然而,计算匹配局部特征点的过程中,存在很多限制因素和要求。如果图像的分辨率不够或无人机的高度太高,则每辆车的像素将相当有限,因此,难以用局部特征点匹配实现目标物的追踪。同时,在无人机拍摄过程中,目标物被建筑物或者树叶遮挡会出现短暂丢失的情况。为了解决上述问题,将光流法结合SIFT运用于车辆追踪,并通过恒速模型融合光流块的模板匹配方法匹配解决追踪丢失问题。

融合SIFT特征点和光流法对运动物体追踪,是将提取的目标物的SIFT特征点导入LK稀疏光流法进行匹配计算,得到目标物特征点的光流向量,从而得到其运动参数,其算法步骤如下<sup>[8]</sup>:

- 1) 在初始帧中选定要追踪的移动目标,提取选定区域目标的SIFT特征点,将其表示为点集  $M = \{(p_i, f_i)\}_{i=1,2,\dots,N}$ ,其中,特征点坐标为  $p_i$ ,SIFT特征描述矢量为  $f_i$ ;
- 2) 依据前一帧追踪目标点集,采用LK算法追踪当前帧中对应的SIFT特征点,用点集  $C = \{(q_i, t_i)\}_{i=1,2,\dots,N}$  表示当前帧中追踪到的SIFT特征点的位置  $q_i, t_i$  对应于  $M$  中的索引值;
- 3) 将当前帧中的全部SIFT特征点表示为  $P = \{(q_i, f_i)\}_{i=1,2,\dots,N}$ ;
- 4) 遍历集合  $P$  和  $M$ ,采用等均值最近邻搜索(ENNS)算法,找到邻近SIFT特征点,得到当前帧中与点集  $M$  对应的集合  $P_1 = \{(q_i, f_i)\}_{i=1,2,\dots,N}$ ,其中,  $P_1$  为

点集  $P$  的子集;

5) 将点集  $C$  和  $P_1$  合并, 并且删除掉重复的 SIFT 特征点, 得运动目标点集  $N = \{(q_i, f_i)\}_{i=1,2,\dots,N}$ ;

6) 判断下一帧是否存在, 不存在则结束, 否则跳转至 2)。

对于突然的加速、减速以及扰动, 提出建立由 SIFT 特征点集合构成的光流块进行光流块匹配的解决方法。在  $t$  时刻的每台车辆形状被标记 9 个点, 纵向和横向各有 5 个点。图 6 中, 蓝色框表示  $t$  时刻处的车辆位置,  $t + \Delta t$  处的目标物光流结果被表示为黑块。当飞行高度在 100 ~ 150 m 之间时, 在纵、横方向上, 如果 1 个光流块覆盖 5 个点中的 3 个以上, 则该块被认为是车辆的一部分, 并且该光流块可以用于追踪此车辆。

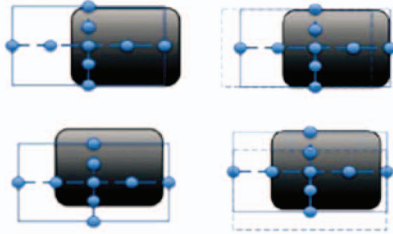


图 6 光流块匹配

Fig. 6 Optical flow block matching

在纵、横方向上, 如果 1 个光流块覆盖特征点数不足 3 个, 则认为车辆被遮挡或者无人机出现大的移动导致目标短暂丢失, 针对该问题, 采用恒速模型融合光流块匹配的策略保持追踪。

物体在相邻两帧的位置  $P_t$  与  $P_{t-1}$  可以用图 7 来表示。

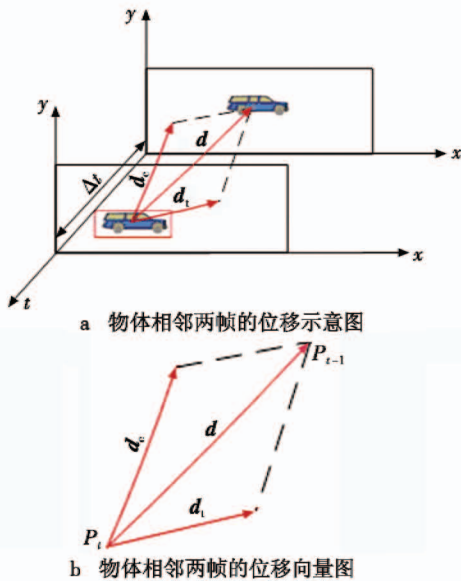


图 7 目标位移图示

Fig. 7 Diagram of target displacement

图中,  $d$  代表两帧间物体的位移, 由  $d_c$  和  $d_t$  决定。

$d_c$  代表仅由在无风情况下摄像机的运动产生的位移, 摄像机运动主要是指由无人机的运动以及自身的震动产生的扰动, 这部分分量对无人机位移的影响较小。 $d_t$  代表由物体的运动引起的位移, 物体的位移主要来自  $d_t$  的贡献。在图像中, 假定车速恒定, 可以通过先验的速度估计当前车的位移, 从而估计新的位置, 即

$$P_t = P_{t-1} + d_t \quad (10)$$

式中,  $d_t = \frac{1}{n} \sum_{i=1}^n d_i$ 。

当目标车辆被遮挡时, 采用恒速模型估计车辆位置, 将估计的车辆位置的光流块与消失时 SIFT 特征点集合构成的光流块进行匹配, 在纵、横方向上, 如果 1 个光流块覆盖 5 个点中的 3 个以上, 则该块被认为是车辆的一部分, 并且该光流块表示车辆的跟踪结果。

### 2.3 追踪结果

将无人机分别悬停在 100 m, 120 m 和 150 m 3 个高度上进行拍摄, 每种情况选取 100 辆车进行追踪, 追踪结果见表 2。

表 2 车辆追踪结果

Table 2 Vehicle tracking results

参数	无人机飞行高度/m		
	100	120	150
追踪时丢失车辆数量	0	1	2
错误率/%	0	1	2

在无人机位于 120 m 的高度拍摄的视频中选取大货车进行追踪, 追踪车辆被红色矩形框框住, 图 8 所示分别为第 1, 4, 7 和 10 帧的追踪结果。

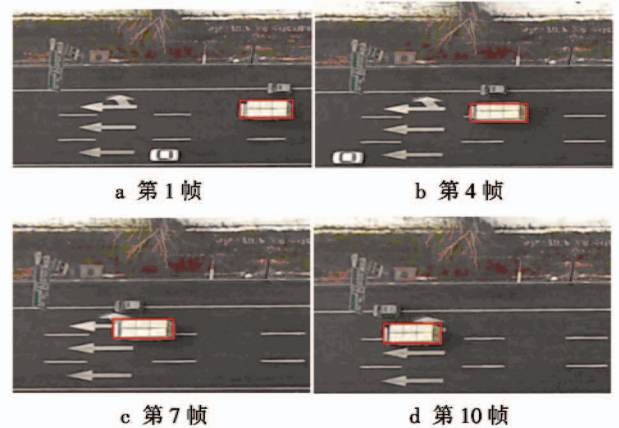


图 8 跟踪结果

Fig. 8 Tracking results

对于干扰的鲁棒性, 在有建筑物遮挡条件下进行目标追踪测试。车辆在未被遮挡前保持正常追踪, 只要发生了遮挡, 预测算法就会对车辆运行进行预测, 当穿过遮挡物时, 利用光流块匹配, 可以持续进行追踪, 实验结果见图 9。

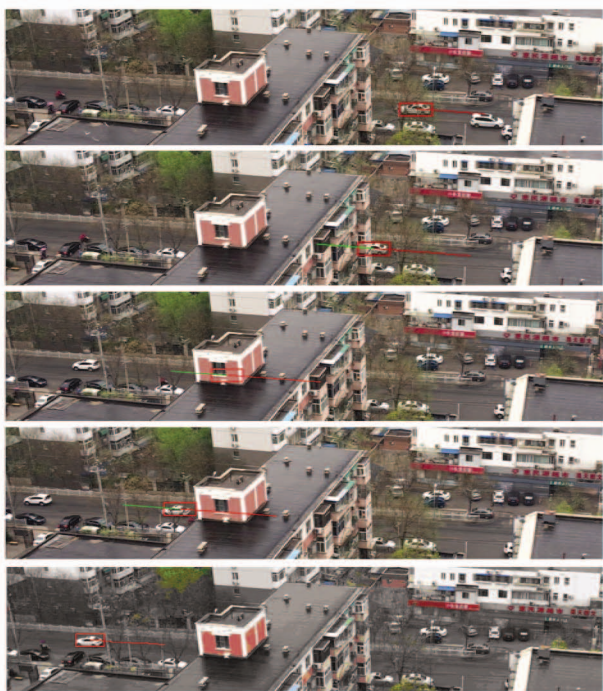


图9 遮挡下的追踪结果

Fig.9 The tracking results when the object is sheltered

基于本文方法的追踪结果表明,追踪效果已经得到了相当大的提高。当无人机飞行高度分别为 100 m, 120 m 和 150 m 时,所有的测试结果都非常可靠,车辆追踪误差不大于 2%。实验结果说明 SIFT 特征匹配融合 LK 追踪算法的方法,不但避免了 LK 追踪算法在计算过程中造成的偏移量误差,同时解决了 SIFT 特征匹配追踪过程中特征点不足的问题,此外,恒速模型融合光流块的模板匹配方法对目标特征点集进行更新和校正,提高了追踪算法的稳定性和鲁棒性。

### 3 小结

本文提出一种采用 3 种图像特征(边缘、光流和局

部特征点)共同作用于无人机视频中的车辆检测和追踪的方法。针对光流法对运动目标检测和追踪出现误检、丢失的问题,应用运动目标纵向的边缘特征配合光流法的方法完成目标检测,应用 SIFT 与光流法实现目标追踪,并提出恒速模型与光流块匹配的方法解决追踪过程中建筑物等对目标的遮挡问题。实验结果表明,本文提出的方法能够对目标进行有效的识别与追踪,并具有较强的鲁棒性。与传统的单图像特征方法相比,本文方法大大提高了追踪的准确性和系统的鲁棒性。

### 参考文献

- [1] 刘亚伟,李小民. 无人机航拍视频中目标检测和跟踪方法综述[J]. 飞航导弹,2016(9):53-56,70.
- [2] 侯阳阳. 基于无人机航拍图像的道路检测[D]. 南京:南京理工大学,2017.
- [3] 朱玮. 基于视觉的四旋翼飞行器目标识别及追踪[D]. 南京:南京航空航天大学,2014.
- [4] 李文辉. 航拍视频中运动目标的检测与追踪算法研究[D]. 西安:西安电子科技大学,2014.
- [5] 刘洁. 基于光流法的运动目标检测和追踪算法研究[D]. 徐州:中国矿业大学,2015.
- [6] KE Y, SUKTHANKAR R. PCA-SIFT: a more distinctive representation for local image descriptors [C]//IEEE Computer Society Conference on Computer Vision & Pattern Recognition,2004:506-513.
- [7] 朱金升. 基于 FPGA 的无人机航拍图像特定目标识别技术应用研究[D]. 济南:山东大学,2017.
- [8] 卢胜男. 复杂交通视频场景中的车辆轨迹提取及行为分析[D]. 西安:长安大学,2016.



请扫描二维码关注我刊