

面向航拍图像多运动目标的实时检测与识别

栾桂芬

(泰州机电高等职业技术学校 信息工程系, 江苏 泰州 225300)

摘要: 为了实现无人机(UAV)航拍图像中多运动目标的实时检测与识别,将静止目标和运动目标分别定义为背景和前景,利用图像稳化技术将航拍图像序列中的每帧与相邻帧对齐,克服UAV飞行动作对摄像机转动拍摄图像的影响;选取图像中的行人和车辆作为前景,分别使用哈尔(Haar-like)特征和级联分类器对图像中的目标进行检测和识别;利用密集光流计算两幅连续图像的运动矢量,从而区分静止目标(背景)和运动目标(前景),最终图像结果仅保留运动目标所在区域;将文章方法用于DroneVehicle航拍数据集实验,每秒平均帧数达到47.08 fps,检测精度为94%,并且表现出较高的召回率和F统计量,可达到实时检测与识别效果。

关键词: 无人机;运动目标;图像稳定;目标检测;密集光流

Real-time Detection and Recognition of Multi-moving Targets in Aerial Images

LUAN Guifen

(Taizhou electromechanical higher vocational and technical school, Department of Information Engineering, Taizhou 225300, China)

Abstract: In order to realize real-time detection and recognition of multiple moving targets in aerial images of unmanned aerial vehicle (UAV), the stationary target and moving target were defined as background and foreground respectively, and the image stabilization technique was used to align each frame in the aerial image sequence with the adjacent frame, so as to overcome the influence of UAV flight movement on the image captured by camera rotation. The pedestrian and vehicle in the image are selected as the foreground, and the haar-like feature and cascade classifier are used to detect and recognize the targets in the image respectively. The dense optical flow is used to calculate the motion vectors of two continuous images, so as to distinguish the stationary target (background) and moving target (foreground), and the final image results only retain the region of the moving target. The proposed method was applied to the DroneVehicle aerial data set experiment, and the average frame per second reached 47.08 FPS, the detection accuracy was 94%, and showed high recall rate and F statistics, which could achieve real-time detection and recognition effect.

Keywords: UAV; Moving target; Image stabilization; Target detection; Dense optical flow

0 引言

无人机(UAV)航拍图像具有分辨率高、视野大以及目标小的特点,携带摄像机的UAV可在任务区域开展实时空中监视^[1]。空中监视可用于行人检测与跟踪、车辆检测与跟踪、人机交互和智能交通系统

等领域^[2]。UAV携带固定摄像机检测运动目标时,通过检索每帧像素中的目标变化对运动目标进行提取和识别^[3]。然而,UAV在实际工况中不可避免的受到环境因素干扰难以检测运动目标。因此,UAV携带非固定摄像机检测运动目标已成为研究重点。

利用非固定摄像机检测运动目标时,通常运用目

收稿日期:2021-11-02; 修回日期:2021-12-08。

作者简介:栾桂芬(1965-),女,江苏泰兴人,硕士,高级讲师,主要从事计算机应用技术、计算机网络技术方向的研究。

引用格式:栾桂芬.面向航拍图像多运动目标的实时检测与识别[J].计算机测量与控制,2022,30(1):221-228.

标分割技术^[4]进行检测。文献 [5] 利用不变矩和动态场景运动分割模型提取 1s 内的帧差实现检测, 但检测速度较慢且无法实时检测。文献 [6] 使用两个或多个连续帧的像素间偏差来识别静止目标和运动目标, 但检测结果依赖于光流计算的准确性和径向畸变量。为了实现非固定摄像机实时检测运动目标, 文献 [7] 通过自适应阈值法消除图像序列背景运动得到前景分割, 结合均值漂移分割方法对检测到的前景进行细化完成运动目标检测。文献 [8] 通过融合时空差异和局部运动轨迹, 利用前景估计分割识别运动目标。然而, 上述方法并没有充分证明非固定摄像机实时检测多运动目标的可靠性。

为了有效地对航拍图像中的多运动目标实时检测与识别, 本文将静止目标和运动目标分别定义为背景和前景, 利用图像稳化技术将航拍图像序列中的每帧与相邻帧对齐, 克服 UAV 飞行动作对摄像机转动拍摄图像的影响。通过加速鲁棒特征 (SURF) 算法检测先前帧和当前帧之间相似的局部特征, 结合仿射变换在齐次坐标中估计的特征位置实现位移估计。使用卡尔曼 (Kalman) 滤波器克服 UAV 飞行动作引起的帧位置变化, 从而对图像中的每帧进行位移补偿。选取图像中的行人和车辆作为运动目标候选 (前景), 分别使用哈尔 (Haar-like) 特征和级联分类器对图像中的目标进行检测和识别, 确定运动目标所在区域并标记前景和背景。利用密集光流计算两幅连续图像的运动矢量, 从而区分静止目标 (背景) 和运动目标 (前景), 最终图像结果仅保留运动目标所在区域。

1 研究准备

1.1 研究视频

本文使用 Visual Studio C++ 在双核 3.40 GHz CPU 和 8 GB RAM 的计算机上处理 UAV 航拍图像。利用天津大学发布的 DroneVehicl 航拍数据集作为研究视频, 数据集记录了大型 UAV 在 100~150 m 的不同飞行高度所采集的图像序列。选取分辨率为 960×540 的三类航拍图像序列 (视频 1、视频 2 和视频 3) 验证本文方法的性能。其中, 视频 1 和视频 2 是由 UAV 在相似高度 (约 110 m) 拍摄以行人和车辆为主要目标的图像序列, 视频 3 是在 150 m 高度拍摄的图像序列, 并且图像中的各目标呈现效果更小。

1.2 研究框架

利用非固定摄像机对多运动目标同时检测需将静止目标 (背景) 和运动目标 (前景) 进行区分, 通过图像稳化技术克服 UAV 飞行动作对摄像机转动拍摄图像影响。UAV 飞行动作对摄像机转动拍摄图像影响的建模, 如图 1 所示。

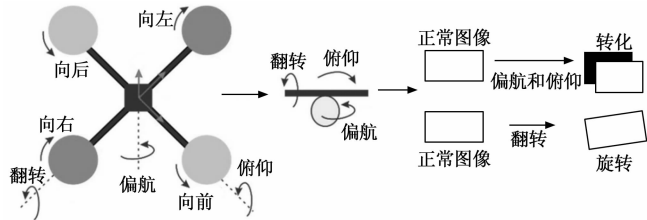


图 1 UAV 飞行动作对摄像机转动拍摄图像影响的建模

非固定摄像机转动相对应的背景运动受 UAV 在偏航、俯仰和翻转三个轴向运动变化的影响。因此, 还需要有效的仿射变换匹配特征点。本文设计的实时运动目标检测和识别系统框架, 如图 2 所示。

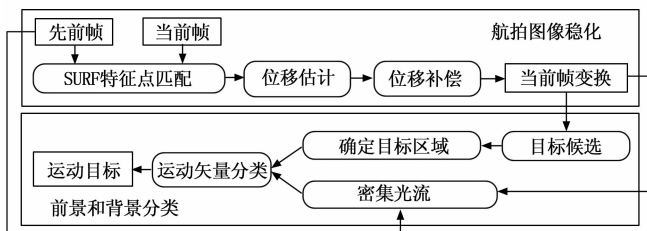


图 2 UAV 实时运动目标检测和识别系统

系统结构又可分为三个主要步骤: (1) 航拍图像稳化; (2) 目标检测与识别; (3) 运动矢量分类。

2 航拍图像稳化

2.1 位移估计

本文使用仿射运动模型处理图像的旋转、缩放和平移, 仿射运动模型可用于估计场景中特定条件下帧间的运动。对于每两个连续帧, 先前帧定义为 $f(t-1)$, 当前帧定义为 $f(t)$ 。为了减少计算时间, 将图像 $f(t)$ 的尺寸缩小到原始尺寸的 75%, 并将颜色设置为灰度模式, 得到新图像 $\hat{f}(t)$ 。使用 SURF 算法^[9]作为特征检测器和描述符找到每个帧上的局部特征。SURF 使用积分图像计算不同的盒式滤波器 (Boxfilter) 来检测图像中的特征点。如果 $f(t)$ 是输入图像, $f_{(x,y)}(t)$ 是 $f(t)$ 在位置 (x,y) 处的像素值, 则图

像 $f(t)$ 的像素值 $P(i, j, t)$ 为:

$$P(i, j, t) = \sum_{x=0}^{x \leq i} \sum_{y=0}^{y \leq j} f_{(x, y)}(t) \quad (1)$$

分别在 x 方向和 y 方向上计算 d_x 和 d_y 对应的 Haar 小波^[10], 围绕每个特征点形成描述符矢量为:

$$v = \left(\sum d_x, \sum d_y, \sum |d_x|, \sum |d_y| \right) \quad (2)$$

以特征点为中心构建 4×4 的数组, 每个向量具有四个方向。因此, 每个特征点总共有 64 个长度矢量。

利用快速近似最近邻查找 (FLANN)^[11] 搜索 $\hat{f}(t-1)$ 和 $\hat{f}(t)$ 之间的特征点对。利用欧氏距离计算所有特征点对的最小距离, 将匹配对确定距离小于 0.6 的特征点对。如果匹配对的总数大于 3 个, 则选择特征点用于下步计算。否则, 使用先前轨迹作为当前运动的估计。

利用仿射变换^[12] 在齐次坐标中估计的特征位置实现位移估计。在齐次坐标中, $\hat{f}(t-1)$ 和 $\hat{f}(t)$ 中的一对特征点之间的关系为:

$$\begin{bmatrix} x(t) \\ y(t) \\ 1 \end{bmatrix} = \mathbf{H} \begin{bmatrix} x(t-1) \\ y(t-1) \\ 1 \end{bmatrix} \quad (3)$$

其中: \mathbf{H} 为齐次仿射矩阵:

$$\mathbf{H} = \begin{bmatrix} 1 + a_{11} & a_{21} & T_x \\ a_{21} & 1 + a_{22} & T_y \end{bmatrix} \quad (4)$$

其中: a_{ij} 为关于旋转角 θ 的参数, T_x 和 T_y 分别为 x 轴和 y 轴上平移 T 的参数。仿射矩阵可以表示为最小二乘问题:

$$\begin{cases} L = mh \\ L = [x(1)' \ y(1)' \ \cdots \ x(\bar{q})' \ y(\bar{q})']^T \\ m = [M_0(1)' \ M_1(1)' \ \cdots \ M_0(\bar{q})' \ M_1(\bar{q})']^T \\ h = [1 + a_{11} \ a_{12} \ T_x \ 1 + a_{21} \ a_{22} \ T_y]^T \end{cases} \quad (5)$$

其中: $q = 1, \dots, \bar{q}$ 为特征的数量顺序,

$$M_0(q)' = (x(q) \ y(q) \ 1 \ 0 \ 0 \ 0) \quad (6)$$

$$M_1(q)' = (0 \ 0 \ 0 \ x(q) \ y(q) \ 1) \quad (7)$$

公式 (5) 中的最优估计 h 通过高斯消元法来最小化均方根误差 (RMSE):

$$RMSE = \frac{1}{Q} \|L - mh'\| = \sqrt{\frac{\sum_{q=1}^Q (L_q - (mh')_q)^2}{Q^2}} \quad (8)$$

由于仿射变换不能表示图像中发生的三维运动, 因此在位移估计中会产生异常值。为了解决这个问题, 在估计过程中采用随机样本一致性 (RANSAC)^[13] 过滤异常值。

2.2 位移补偿

使用 Kalman 滤波器^[14] 对平移和旋转轨迹进行位移补偿, 从而为每帧生成新的变换。Kalman 滤波器由预测和测量校正两部分组成。预测部分估计轨迹 $\hat{z}(t) = [T_x(t), T_y(t), \hat{\theta}(t)]$ 在 $\hat{f}(t)$ 处的状态为:

$$\hat{z}(t) = \hat{z}(t-1) \quad (9)$$

其中, 初始状态由 $z(0) = [0, 0, 0]$ 定义, 误差的协方差可以估计为:

$$\hat{e}(t) = e(t-1) + \Omega_p \quad (10)$$

其中: 初始误差的协方差由 $e(0) = [1, 1, 1]$ 定义, Ω_p 为过程噪声的协方差。最佳 Kalman 增益计算为:

$$K(t) = \frac{\hat{e}(t)}{e(t) + \Omega_m} \quad (11)$$

其中: Ω_m 为测量噪声的协方差。误差的协方差可通过以下方式进行补偿:

$$e(t) = (1 - K(t))\hat{e}(t) \quad (12)$$

测量校正补偿 $\hat{f}(t)$ 处的轨迹状态, 可计算为:

$$z(t) = z(t) + K(t)(\Gamma(t) - z(t)) \quad (13)$$

其中: 新状态包含由 $z(t) = [T'_x(t), T'_y(t), \theta'(t)]$ 定义的补偿轨迹, $\Gamma(t)$ 为轨迹测量的累加:

$$\begin{aligned} \Gamma(t) = \sum_{\tau=1}^{t-1} [(\bar{T}_x(\tau) + T_x(t), (\bar{T}_y(\tau) + T_y(t)), \\ (\bar{\theta}(\tau) + \theta(t))] = [\Gamma_x(t), \Gamma_y(t), \Gamma_\theta(t)] \end{aligned} \quad (14)$$

因此, 得到的新轨迹为:

$$\begin{aligned} [\bar{T}_x(t), \bar{T}_y(t), \bar{\theta}(t)] = [T_x(t), T_y(t), \theta(t)] + \\ [\sigma_x(t), \sigma_y(t), \sigma_\theta(t)] \end{aligned} \quad (15)$$

其中: $\sigma_\kappa(t) = \bar{T}_\kappa(t) - \Gamma_\kappa(t), \kappa = x, y, \theta$ 。

在新的图像平面中, 利用新轨迹得到当前帧中的变换 $\bar{f}(t)$:

$$\begin{aligned} \bar{f}(t) = \\ f(t) \begin{bmatrix} \Phi(t) \cos \bar{\theta}(t) & -\Phi(t) \sin \bar{\theta}(t) \\ \Phi(t) \sin \bar{\theta}(t) & \Phi(t) \cos \bar{\theta}(t) \end{bmatrix} + \begin{bmatrix} \bar{T}_x(t) \\ \bar{T}_y(t) \end{bmatrix} \end{aligned} \quad (16)$$

其中: $\Phi(t)$ 为比例因子:

$$\Phi(t) = \frac{\cos \bar{\theta}(t)}{\cos \left(\tan^{-1} \left(\frac{\sin \bar{\theta}(t)}{\cos \bar{\theta}(t)} \right) \right) (t)} \quad (17)$$

3 目标检测与识别

航拍图像的每帧通过图像稳化技术确定背景和前景,其中,前景由图像中的候选运动目标(行人和车辆)组成。本文分别使用 Haar-like 特征^[15]和增强级联分类器^[16]对前景进行检测和识别。

3.1 Haar-like 特征

Haar-like 特征主要检测图像中不同尺寸的目标。Haar-like 特征的模板,如图 3 所示。

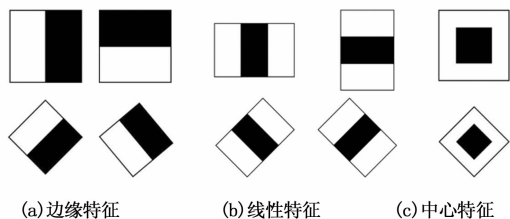


图 3 Haar-like 特征的模板

每个特征模板内有白色和黑色两种矩形组组成,白色组和黑色组中的像素强度值分别累积,相邻组之间的区别给出了亮区和暗区。因此,Haar-like 特征适合于定义图像中的信息,从而寻找不同尺寸的目标。

Haar-like 特征值计算为像素灰度值在黑色矩形和整个特征区域上的加权和,使用积分图^[17]最小化图像矩形区域像素总和中的数组。本文所采用的航拍图像示例,如图 4 所示。其中,图 4 (a)和图 4 (b)为本文要研究的主要运动目标示例,图 4 (c)为非运动目标示例,即路标、栅栏、道路图案、草地、电线、路障等。

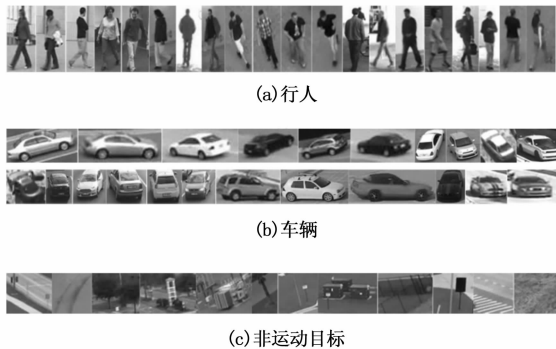


图 4 图像示例

3.2 级联分类器

使用 AdaBoost 算法^[18]对所选分类器的特征进行组合,选择分类器作为阈值来确定每个特征的最佳分

类函数。训练样本设置为 $(\alpha_s, \beta_s), s = 1, 2, \dots, N$,其中, α_s 表示样本的类标签, $\beta_s = 0$ 或 1 分别表示负标签或正标签。将每个样本转换为灰度,然后按比例缩小到检测器的基准分辨率。AdaBoost 算法为迭代中的所有训练样本创建权重向量。所有样本 $(\alpha_1, \beta_1), \dots, (\alpha_N, \beta_N)$ 的初始权重向量设置为 $\omega_1(s) = 1/N$,则与所选分类器相关的误差为:

$$\epsilon_i = \sum_{s=1}^N \omega_i(s) \cdot \lambda_i(\alpha_s) \neq \beta_s \quad (18)$$

$\lambda_i(\alpha_s) = 0$ 或 1 分别是负标签或正标签的选定分类器, $i = 1, \dots, I$ 为迭代数量。所选分类器将权重向量更新为:

$$\omega_{i+1}(s) = \omega_i(s) \delta_i^{1-r_s} \quad (19)$$

其中:当类标签 α_s 分类正确时,则 $r_s = 0$;否则, $r_s = 1$ 。 δ_i 为权重参数:

$$\delta_i = \frac{\epsilon_i}{1 - \epsilon_i} \quad (20)$$

最终的分类器 $W(\alpha)$ 是每个区域的标记结果:

$$W(\alpha) = \begin{cases} 1, & \text{若 } \sum_{i=1}^I \left[\log\left(\frac{1}{\delta_i}\right) \times \lambda_i(\alpha) \right] \geq \frac{1}{2} \sum_{i=1}^I \log\left(\frac{1}{\delta_i}\right) \\ 0, & \text{其他} \end{cases} \quad (21)$$

用于目标检测与识别的级联分类器结构,如图 5 所示。

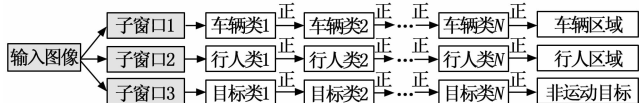


图 5 目标检测与识别的级联分类器

图 5 中的子窗口在图像上滑动来识别包含目标的区域,区域在每个分类器被标记为正或负。如果区域标记为正,则分类器将进入下一阶段,即该区域识别目标。否则,该区域将标记为负并消除。最后显示处运动目标候选区域,而非运动目标区域不显示在图像中。如果运动目标候选区域与非运动目标相同,则该区域作为前景被消除。令第 n 个前景区域表示为:

$$Obj[n] = [(x_{\min}(n), y_{\min}(n)), (x_{\max}(n), y_{\max}(n))] \quad (22)$$

其中: $(x_{\min}(n), y_{\min}(n))$ 和 $(x_{\max}(n), y_{\max}(n))$ 分别是矩形前景像素位置的最小位置和最大位置。通过

与非运动目标区域的比较, 可以消除运动目标候选区域的误检, 从而减少计算时间。

4 运动矢量分类

利用密集光流计算两幅连续图像的运动矢量, 使得静止目标(背景)和运动目标(前景)之间的运动方向更加清晰便于区别。运动方向定义为背景和前景之间高度重复运动矢量的值。其中, 背景的运动方向受摄像机转动的影响。如果前景与背景具有相同的运动方向, 则目标从前景中忽略。因此, 最终图像结果仅保留运动目标所在区域。

采用 Farneback 光流法^[19]得到两幅连续图像的运动矢量。假设在 10×10 像素的窗口 $G(j)$ 内选择像素 j , 通过使用多项式展开, $G(j)$ 中的每个像素可以由 $f(t-1)$ 处的多项式来近似:

$$f_{ks}^p(t-1) = \mathbf{p}^T \mathbf{A}(t-1) \mathbf{p} + \mathbf{b}^T(t-1) \mathbf{p} + c(t-1) \quad (23)$$

其中: \mathbf{p} 为向量, $\mathbf{A}(t-1)$ 为对称矩阵, $\mathbf{b}(t-1)$ 为向量, $c(t-1)$ 为标量。 $f(t)$ 处的局部坐标系可以定义为:

$$f_{ks}^p(t) = \mathbf{p}^T \mathbf{A}(t) \mathbf{p} + \mathbf{b}^T(t) \mathbf{p} + c(t) \quad (24)$$

在 $f(t)$ 处用全局位移 Δt 为 $f_{ks}^p(t)$ 构造新信号, 两个输入图像的局部坐标系之间的关系为:

$$f_{ks}^p(t) = ((\mathbf{p} - \Delta t))^T \mathbf{A}(t-1) ((\mathbf{p} - \Delta t)) + \mathbf{b}^T(t-1) (\mathbf{p} - \Delta t) + c(t-1) = \mathbf{p}^T \mathbf{A}(t-1) \mathbf{p} + (\mathbf{b}(t-1) - 2\mathbf{A}(t-1)\Delta t)^T \mathbf{p} + \Delta^T \mathbf{t} \mathbf{A}(t-1) \Delta t - \mathbf{b}^T(t-1) \Delta t - c(t-1) \quad (25)$$

公式 (24) 和公式 (25) 的系数可以等效为:

$$\begin{cases} \mathbf{A}(t) = \mathbf{A}(t-1) \\ \mathbf{b}(t) = \mathbf{b}(t-1) - 2\mathbf{A}(t-1)\Delta t \\ c(t) = \Delta t (\Delta^T \mathbf{t} \mathbf{A}(t-1) - \mathbf{b}^T(t-1)) + c(t-1) \end{cases} \quad (26)$$

因此, 运动目标所在区域中提取的总位移为:

$$\Delta t = -\frac{1}{2} \mathbf{A}^{-1}(t-1) (\mathbf{b}(t) - \mathbf{b}(t-1)) \quad (27)$$

公式 (27) 中的位移是由 x 轴 ($\Delta_x t$) 和 y 轴 ($\Delta_y t$) 组成的每个对应运动目标所在区域的平移, 因此, 运动矢量的角度值为:

$$\Delta_{\theta} t = \tan^{-1} \left(\frac{\Delta_{(x+1,y)} t - \Delta_{(x-1,y)} t}{\Delta_{(x,y+1)} t - \Delta_{(x,y-1)} t} \right) \times \frac{180}{\pi} \quad (28)$$

由于每个 10×10 像素的邻域计算运动矢量, 总位移为 (图像宽度/10) \times (图像高度/10) 的矩阵。

因此, 新的第 n 个前景区域为:

$$Obj2[n] = \left[\frac{(x_{\min}(n), y_{\min}(n))}{10}, \frac{(x_{\max}(n), y_{\max}(n))}{10} \right] \quad (29)$$

利用 Farneback 光流估计的示例, 如图 6 所示。其中, 图 6 (a) 的运动目标所在区域分别标记行人和车辆的运动目标候选(前景), 图 6 (b) 为估计的运动矢量分布示例。

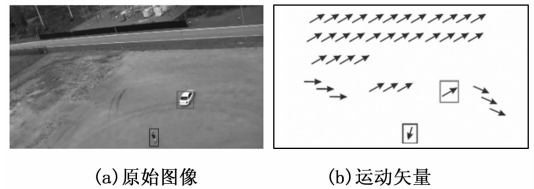


图 6 Farneback 光流估计

在固定摄像机航拍图像中, 背景中的运动矢量为零, 即运动方向值为零, 则两个连续帧之间没有移动(由箭头方向表示)。在本文的示例中, 图像由非固定摄像机拍摄, 背景中的运动矢量有不同的方向, 如图 6 (b) 所示。上面的运动目标所在区域是分类为非运动目标的停放车辆, 其运动矢量与背景中的大多数运动矢量相似。下面的运动目标所在区域是分类为运动目标的行人, 其运动向量与背景中的大多数运动矢量不同。因此, 每个运动目标候选上的运动方向作为每个的运动目标所在区域中出现次数最多的运动矢量。在背景中, 运动方向可以作为除前景外图像中出现次数最多的运动矢量。运动矢量分类和运动目标选择的流程图, 如图 7 所示。

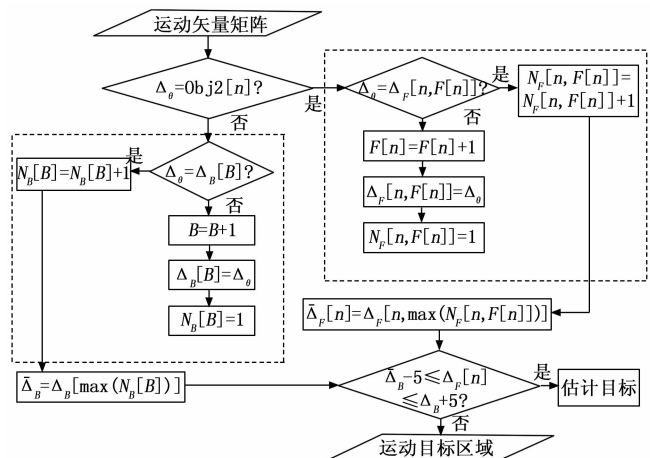


图 7 运动矢量分类和运动目标选择的流程图

在每个运动目标所在区域中，角度值等于或大于零的运动矢量划分为同类。如果运动矢量在背景中，则将其分类为 $\Delta_B[B]$ ，其中， B 为背景中类别的数量顺序，且每个类别的总数记为 $N_B[B]$ 。如果运动矢量在前景中，则将其分类为 $\Delta_F[n, F(n)]$ ，其中， $F(n)$ 为前景中类别的数量顺序，且每个类别的总数记为 $N_F[n, F(n)]$ 。将背景 $\bar{\Delta}_B$ 和第 n 前景 $\bar{\Delta}_F[n]$ 的运动方向分别确定为最大 Δ_B 和 $\Delta_F[n]$ ，如果 $\bar{\Delta}_F[n]$ 的值属于 $\bar{\Delta}_B$ 的阈值，则该目标识别为非运动目标。否则，该目标识别为运动目标。图像将仅显示运动目标的所在区域。文献 [20] 指出背景和静态目标之间的运动方向差异较小，不超出阈值范围 $[-5, +5]$ 。因此，背景中运动方向的最小阈值和最大阈值分别为 -5 和 $+5$ 。

5 实验分析

5.1 运动矢量的检测结果

UAV 携带非固定摄像机的航拍图像具有不稳定性，从而导致图像关于静态（非运动）目标和动态（运动）目标的运动矢量不适合区分与辨别。无图像稳化和有图像稳化的运动矢量结果，分别如图 8 和图 9 所示。其中，图 8 (a) 和图 9 (a) 给出了背景中的运动矢量。图 8 (b) 和图 9 (b) 给出了作为静态目标（停放车辆）的运动矢量。图 8 (c) 和图 9 (c) 给出了作为动态目标（行人）的运动矢量。

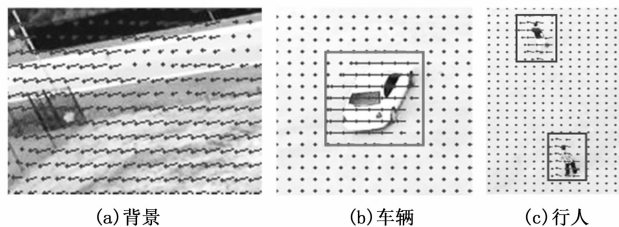


图 8 无图像稳化的运动矢量结果

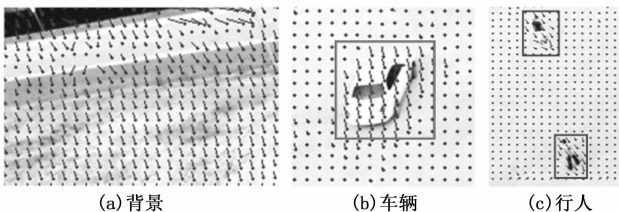


图 9 有图像稳化的运动矢量结果

从图 8 可以看出，动态目标和静态目标的运动矢量基本相同，但与背景中的运动矢量略有不同。因此，无图像稳化的运动矢量结果不准确。图 9 (b) 中，停放车辆（静态目标）的运动矢量与背景中的运动矢量基本相同。图 9 (c) 中，行人（动态目标）的运动矢量与背景中的运动矢量有很大不同。因此，经过图像稳化后的运动矢量结果适用于区分静态目标和动态目标。

5.2 运动目标的检测结果

视频 1、视频 2 和视频 3 的运动目标的检测和识别结果，分别如图 10、图 11 和图 12 所示。

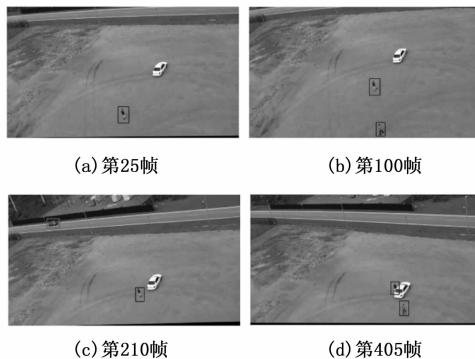


图 10 视频 1 中的运动目标检测结果

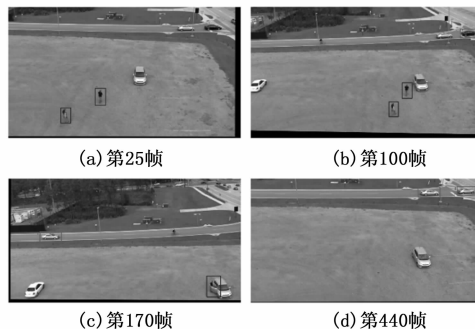


图 11 视频 2 中的运动目标检测结果

由图 10 至图 12 可见，本文方法利用运动矢量的分类结果将区域未被识别为前景的目标进行了消除，从而对多运动目标进行检测与识别。在图 12 中包含 5 个行人，而图 12 (a) 和图 12 (c) 仅检测到 4 个行人。因此，如果目标只有微小的位移，则很难区分运动矢量，而最终导致误检为非运动目标。

5.3 检测性能分析

本文方法对 DroneVehicl 航拍数据集中视频 1 至视频 3 的计算时间性能结果，如表 1 所示。

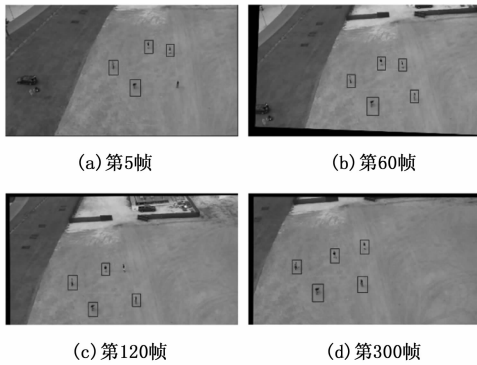


图 12 视频 3 中的运动目标检测结果

表 1 计算时间性能

| 视频 | 每秒平均帧数/fps |
|------|------------|
| 视频 1 | 49.9 |
| 视频 2 | 42.16 |
| 视频 3 | 49.2 |
| 平均值 | 47.08 |

本文利用真阳性 (TP)、假阳性 (FP)、假阴性 (FN)、精度 (PR)、召回率 (R) 和 F 统计量 (F 值) 作为检测性能的指标。其中, TP 是与运动目标对应的检测区域。 FP 是检测到的与运动目标无关的区域。 FN 是与未检测到的运动目标相关联的区域。因此, 精度 (PR)、召回率 (R) 和 F 统计量 (F 值) 可以分别计算为:

$$PR = \frac{TP}{TP + FP} \quad (30)$$

$$R = \frac{TP}{TP + FN} \quad (31)$$

$$F = 2 \times \frac{PR \times R}{PR + R} \quad (32)$$

本文方法对 DroneVehicl 航拍数据集中视频 1 至视频 3 的检测性能结果, 如表 2 所示。

表 2 检测性能结果

| 视频 | TP | FP | FN | PR | R | F 值 |
|------|------|------|------|------|------|-------|
| 视频 1 | 124 | 7 | 6 | 0.95 | 0.95 | 0.95 |
| 视频 2 | 245 | 19 | 23 | 0.92 | 0.91 | 0.91 |
| 视频 3 | 184 | 12 | 25 | 0.94 | 0.88 | 0.90 |
| 平均值 | | | | 0.94 | 0.91 | 0.92 |

由表 2 可见, 视频 1 具有最高的精度和召回率, 视频 2 精度最低, 视频 3 召回率最低。这是由于视频

1 中只有少量目标且目标尺寸相对较大, 视频 2 中存在与被检测目标 (行人和车辆) 相似的目标, 例如树木、栅栏、路标、房屋和灌木丛, 而视频 3 中的目标移位较小。

将本文方法与现有文献方法对 DroneVehicl 航拍数据集中视频 1 至视频 3 的检测进行了比较, 其他方法包括: 动态运动分割^[7]、轨迹分类^[8]、计算机视觉^[21]、多图匹配^[22]和区域生长算法^[23]。本文方法与其他方法的性能结果对比, 如表 3 所示。

表 3 本文方法与其他方法的性能结果对比

| 方法 | 每秒平均帧数/fps | PR | R | F 值 |
|------------------------|------------|------|------|-------|
| 本文方法 | 47.08 | 0.94 | 0.91 | 0.92 |
| 动态运动分割 ^[7] | — | 0.66 | 0.86 | 0.74 |
| 轨迹分类 ^[8] | 1 | 0.70 | 0.76 | 0.72 |
| 计算机视觉 ^[21] | — | 0.94 | 0.89 | 0.91 |
| 多图匹配 ^[22] | 1.6 | — | — | 0.73 |
| 区域生长算法 ^[23] | 5 | — | — | 0.76 |

由表 3 可见, 本文方法在多运动实时检测与识别中的性能最佳, 每秒平均帧数达到 47.08 fps, 比其他方法快, 并且平均精度和召回率分别为 0.94 和 0.91。实验结果表明, 本文方法通过 Farneback 光流计算两个连续帧之间的运动矢量场, 并且具有快速的计算时间。进一步证明了利用运动矢量分类的重要性, 这是由于运动矢量分类可对静止目标和运动目标进行区分, 确定背景和前景中的运动方向, 从而有效的检测并识别 UAV 航拍图像中的多个运动目标。

6 结束语

本文提出了一种面向携带非固定摄像机 UAV 的航拍图像多运动目标实时检测与识别方法。通过图像稳化技术将航拍图像序列中的每帧与相邻帧对齐, 克服 UAV 飞行动作对摄像机转动拍摄图像的影响。选取图像中的行人和车辆作为前景, 分别使用哈尔 (Haar-like) 特征和级联分类器对图像中的目标进行检测和识别。利用密集光流计算两幅连续图像的运动矢量, 结合运动矢量分类方法消除运动方向与背景相似的运动目标候选, 从而区分静止目标 (背景) 和运动目标 (前景), 最终图像结果仅保留运动目标所在区域。通过对不同航拍图像序列的结果进行比较,

验证了本文方法在复杂背景下具有计算速度快和精度高的特点。

由于多运动目标实时检测与识别是从航拍图像稳态开始, 如果目标位移较小, 则 SURF 算法检测局部特征变化较小。因此, 本文方法的局限性在于检测精度取决于目标在图像中的位移变化。在未来的研究中, 利用更有效的图像稳态技术逐步提升目标位移估计。

参考文献:

- [1] 王 丹. 无人机航拍技术及其应用 [J]. 电子制作, 2020 (20): 54-55, 49.
- [2] 杨岳航, 陈武雄, 朱 明, 等. 基于机器视觉的无人机自主着陆技术 [J]. 国外电子测量技术, 2020, 39 (4): 57-61.
- [3] 江 波, 屈若锟, 李彦冬, 等. 基于深度学习的无人机航拍目标检测研究综述 [J]. 航空学报, 2021, 42 (4): 137-151.
- [4] 卢 笑, 曹意宏, 周炫余, 等. 基于深度强化学习的两阶段显著性目标检测 [J]. 电子测量与仪器学报, 2021, 35 (6): 34-42.
- [5] TAN Li, LÜ Xinyue, LIAN Xiaofeng, et al. YOLOv4 _ Drone: UAV image target detection based on an improved YOLOv4 algorithm [J]. Engineering, 2021, 93 (1): 107-121.
- [6] 郑 锴, 郑献民, 殷少锋, 等. 基于最小二乘拟合的无人机运动目标测速方法 [J]. 兵器装备工程学报, 2021, 42 (3): 61-64.
- [7] 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.
- [8] JULIUS I, MICHEL K, LUC M, et al. Bounded-error target localization and tracking using a fleet of UAVs [J]. Automatica, 2021, 132 (2): 109-115.
- [9] 罗久飞, 邱 广, 张 毅, 等. 基于自适应双阈值的 SURF 双目视觉匹配算法研究 [J]. 仪器仪表学报, 2020, 41 (3): 240-247.
- [10] 李水泉, 乐 阳, 李清泉, 等. 一种 Haar 小波概要的流数据压缩方法 [J]. 武汉大学学报 (信息科学版), 2021, 46 (8): 1216-1223.
- [11] 徐 明, 刁 燕. 基于 SURF 算子与 FLANN 搜索的图像匹配方法研究 [J]. 现代计算机, 2020 (14): 49-52, 57.
- [12] 马雪梅, 刘森晶. 基于仿射变换的分镜头动画图像合成技术研究 [J]. 电子设计工程, 2020, 28 (24): 178-181, 187.
- [13] 丁 一. 基于 ORB 和 RANSAC 算法的图像匹配查重复算法 [J]. 电子技术与软件工程, 2021 (11): 134-135.
- [14] 叶春煦, 陶德桂. 无人旋翼机动平台降落轨迹估计方法研究 [J]. 电子测量技术, 2020, 43 (16): 132-136.
- [15] 孙凯传, 柳晨华, 姚光顺, 等. 联合最小软阈值二乘和 Haar-like 特征匹配的视觉跟踪 [J]. 激光与光电子学进展, 2019, 56 (24): 85-91.
- [16] 金东勇, 陈俊霞. 车载视频交通场景定位与文字识别方法的改进 [J]. 单片机与嵌入式系统应用, 2018, 18 (10): 55-58.
- [17] 易三莉, 张桂芳, 贺建峰, 等. 基于最大类间方差的最大熵图像分割 [J]. 计算机工程与科学, 2018, 40 (10): 1874-1881.
- [18] WANG Wenyang, SUN Dongchu. The improved AdaBoost algorithms for imbalanced data classification [J]. Information Sciences, 2021, 563 (1): 358-374.
- [19] 郑萌萌, 钱慧芳, 周 璇. 基于监控视频的 Farneback 光流算法的人体异常行为检测 [J]. 国外电子测量技术, 2021, 40 (3): 16-22.
- [20] DAISY J P, THOMAS J M, EDWIN D. Motion direction tuning in centre-surround suppression of contrast [J]. Vision Research, 2021, 179 (2): 85-93.
- [21] 刘 芳, 吴志威, 杨安喆, 等. 基于多尺度特征融合的自适应无人机目标检测 [J]. 光学学报, 2020, 40 (10): 133-142.
- [22] BAI X, YANG M J. UAV based accurate displacement monitoring through automatic filtering out its camera's translations and rotations [J]. Journal of Building Engineering, 2021, 44 (3): 102-109.
- [23] ZHOU Y, RUI T, LI Y R, et al. A UAV patrol system using panoramic stitching and object detection [J]. Computers and Electrical Engineering, 2019, 80 (4): 106-123.