

DOI: 10.19650/j.cnki.cjsi.J1905657

抖动干扰下运动目标精准检测与跟踪算法设计^{*}

郑 浦, 白宏阳, 李政茂, 郭宏伟

(南京理工大学能源与动力工程学院 南京 210094)

摘 要:针对在抖动干扰下运动目标检测精度较差的问题,提出了一种基于光流法与三帧差分法的运动目标检测算法,首先用基于卢卡斯-卡那得(LK)光流法的稳像算法对视频去抖,然后用三帧差分法提取目标。仿真结果表明,稳像后的峰值信噪比(PSNR)值提高了3.6 dB左右,所设计算法在抖动干扰下能够准确提取出目标,在测试平台上的平均处理速度为28 fps;同时,针对传统核相关滤波(KCF)算法对尺度变化和部分遮挡目标跟踪性能较差的问题,设计了一种改进的KCF算法,通过对目标构造图像金字塔,计算滤波器在图像金字塔不同层上的响应,找到响应最大层并更新下一帧目标位置,同时加入了遮挡检测机制,减小目标遮挡对跟踪的影响。仿真结果表明,改进后的算法对尺度变化和部分遮挡的目标跟踪鲁棒性更优,可实现对目标的稳定跟踪,处理速度为33 fps。通过与KCF算法进行比较说明该算法的准确率提高了4.2%,成功率提高了11.8%。

关键词:电子稳像;Lucas-Kanade光流法;目标检测;核相关滤波器;目标跟踪

中图分类号: TH89 **文献标识码:** A **国家标准学科分类代码:** 460.4099

Design of accurate detection and tracking algorithm for moving target under jitter interference

Zheng Pu, Bai Hongyang, Li Zhengmao, Guo Hongwei

(College of Energy and Power Engineering, Nanjing University of Science and Technology, Nanjing 210094, China)

Abstract: Aiming at the problem of poor detection accuracy for moving targets under jitter interference, a moving target detection algorithm based on optical flow method and three-frame difference method is proposed. Firstly the image stabilization algorithm based on LK (Lucas-Kanade) optical flow method is used to de-jitter the video, then the three-frame difference method is used to extract the target. Simulation results show that the PSNR (Peak Signal to Noise Ratio) value is increased by 3.6 dB after image stabilization, and the designed algorithm can accurately extract the target under jitter interference, the average processing speed on the test platform is 28 fps. At the same time, aiming at the problem that traditional KCF (Kernelized Correlation Filter) algorithm has poor tracking performance for scale-changing and partially occluded targets, an improved KCF algorithm is designed, which constructs the image pyramid of the target, then calculates the filter response on different layers of the image pyramid, finds the layer with largest response and updates the target location of next frame. Meanwhile, an occlusion detection mechanism is introduced in the algorithm, which reduces the impact of target occlusion on tracking. Simulation results indicate that the improved algorithm has stronger robustness to the scale-changing and partially occluded targets, and can achieve stable target tracking, the processing speed of the algorithm is 33 fps. Compared with the KCF algorithm, the precision of the proposed algorithm is increased by 4.2% and the success rate is increased by 11.8%.

Keywords: electronic image stabilization; Lucas-Kanade optical flow method; object detection; kernelized correlation filter (KCF); object tracking

0 引言

运动目标检测与跟踪是计算机视觉领域的重要研究方向之一,在公共安全、道路交通、视频监控、军事侦察、精确制导等军用和民用领域均有着重要的作用^[1]。目前,目标检测和跟踪仍然面临着许多问题,在检测过程中,目标可能会发生一些变化,比如姿态或形状的变化、尺度的变化、光线亮度的变化以及背景的变化、相机的随机抖动等^[2],从而影响目标检测和跟踪的精度。

针对目标检测问题,国内外学者作了大量研究。帧间差分法适合静态背景中的目标提取,速度较快,但是目标提取不够完整,且容易出现“重影”^[3]。背景减法法通过当前帧与背景差分提取目标,该方法的关键在于建立背景模型,越是复杂的背景模型,其检测效果越好,但是时间复杂度和空间复杂度较高,因此如何减少背景建模算法的计算量,提高算法实时性是一个很重要的问题^[4-5]。光流法既能实现静态背景下的目标检测,也可实现动态背景下运动目标的检测,但是光流法对噪声比较敏感,计算量也较大^[6]。基于背景减除、帧间差分和光流的算法,对噪声都比较敏感,当视频序列受到抖动干扰时,这些方法的准确性将大幅度降低,针对上述情况,本文设计了一种光流法与三帧差分法结合的目标检测算法和思路,对于输入的包含目标的抖动视频序列,首先采用基于光流的电子稳像算法对抖动视频进行去抖处理,然后将去抖后的视频输入到三帧差分法模块中,采用三帧差分法提取前景目标。

从数学模型的角度出发,目标跟踪算法可分为如下两类:生成式模型和判别式模型^[7]。生成式跟踪模型通过学习,得到目标的跟踪模板,然后根据模板,通过一定的匹配准则,在下一帧图像的搜索区域内进行模式匹配,从而完成对目标的跟踪。生成式模型由于学习和计算过程复杂,实时性比较差,典型的生成式算法有马尔科夫模型、贝叶斯网络模型^[8]等。Tuzel 等^[9]通过融合颜色特征和颜色协同特征,并加入到贝叶斯算法中,使得算法能够在光线变化的环境下工作,提高了跟踪算法的鲁棒性。吴孟俊等^[10]基于动态贝叶斯网络建立状态模型,并利用粒子滤波方法对颜色特征和梯度特征进行融合,较好地解决复杂环境下的目标跟踪问题。判别式跟踪模型把跟踪问题看作是二分类问题,通过图像序列将获得的样本送入分类器进行训练,得到能够区分背景和前景的分类器。典型的判别式算法有核相关滤波(kernelized correlation filters, KCF)跟踪算法、TLD(tracking-learning-detection, TLD)、Struck 等^[11-13],其中 KCF 无论是速度还是精度,都有着很好的表现。但是原始的 KCF 算法输入特征为灰度图像,这种特征较为简单,使得算法易受光照

等外界因素的影响,跟踪效果不太理想。近年来有大量学者对 KCF 算法进行了研究,在算法中加入了更加有效、鲁棒的特征,提高了算法的跟踪效果。Dalal 等^[14]采用梯度直方图特征,将图像的边缘梯度信息融合到 KCF 算法中,有效地提升了跟踪效果。Li 等^[15]将图像的灰度信息、颜色信息以及 HOG 特征融合,得到了更为鲁棒的跟踪结果。Zhu 等^[16]将图像的梯度信息与颜色信息融合,有效提升了算法的性能。

本文针对 KCF 算法对于尺度变化、部分遮挡目标跟踪鲁棒性较差的问题,设计了一种改进的 KCF 跟踪算法,改进后的算法对于尺度变化和部分遮挡目标的跟踪鲁棒性明显提高。

1 目标检测算法设计

1.1 基于光流的电子稳像方法

在目标检测过程中,相机可能会出现随机抖动,严重干扰了检测结果。如果抖动幅度较小,可以近似地将抖动看成是仿射变换,若求得相邻帧之间的仿射变换参数,便可以对抖动进行补偿,因此视频去抖的关键就在于准确估计全局运动参数。目前国内外估计全局运动参数的方法主要有块匹配法、灰度投影法以及特征跟踪法等^[17],其中块匹配法对块的大小选择非常敏感,灰度投影法在灰度信息单一的场景中处理效果较差,并且这两种方法均是用于平移运动的估计,而特征跟踪法对平移、旋转、放缩等均能够很好地估计。因此,本文提出一种基于光流的特征跟踪电子稳像算法,以消除相机的微弱抖动对目标检测精度的影响。所设计的电子稳像算法流程如图 1 所示,该算法主要包括:1)提取视频中的相邻两帧图像,并进行预处理,然后采用 SURF 算法检测前一帧图像中的特征点;2)采用 LK(Lucas-Kanade)光流法预测下一帧中对应的特征点;3)采用随机采样一致(random sample consensus, RANSAC)算法剔除上下两帧图像中误匹配的特征点;4)根据特征点计算两帧图像之间的仿射变换参数;5)采用卡尔曼滤波对计算得到的仿射变换参数进行平滑处理,得到平滑的全局运动矢量;6)对后一帧图像进行仿射变换,得到稳定的视频序列。

目前使用较多的特征点提取方法有尺度不变特征变换算法(scale-invariant feature transform, SIFT)、加速鲁棒特征算法(speeded up robust features, SURF)以及定向二进制简单描述符(oriented fast and rotated brief, ORB)3 种,其中 ORB 算法计算速度最快,但是旋转和尺度的鲁棒性都不如 SURF;SIFT 算法准确度最好,但是计算速度难以达到实时性要求,综合考虑算法速度、准确率等要素,本文采用 SURF 算法提取特征点。SURF 特征点是由 Herbert 等在 2006 年提出的一种稳健的局部特征点检测

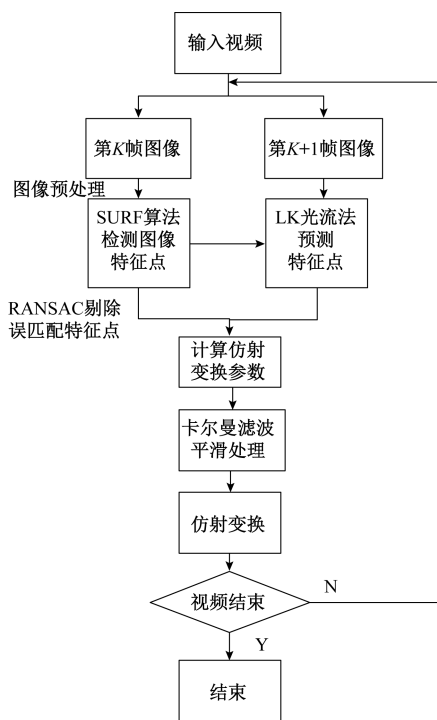


图1 电子稳像算法流程

Fig.1 The flow chart of the electronic image stabilization algorithm

分后的图像进行逻辑与操作,可消除双影问题,并抑制噪声。

三帧差分法流程如图2所示。首先从视频中提取相邻的三帧图像,其中相邻两帧两两进行灰度化和差分操作,得到两张差分后的图像,再进一步对这两张差分图像进行与运算和二值化操作,得到三帧差分图像。针对三帧差分法容易出现“空洞”的现象,对差分后的图像进行形态学膨胀操作,然后寻找膨胀后图像的轮廓和最大连通域,根据最大连通域得到目标的最小外接矩形。

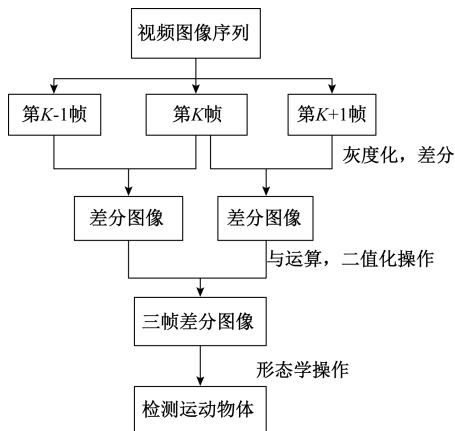


图2 三帧差分算法流程

Fig.2 The flow chart of three-frame difference algorithm

和描述算法^[18]。它是 SIFT 算法的改进, SURF 使用海森矩阵作特征点检测并用积分图加速运算,同时对特征描述子进行了降维处理,从而提升了算法的执行效率, SURF 特征点具有良好的旋转、尺度鲁棒性。

确定特征点之后,用 LK 光流法预测其在下一帧的位置,由于光流法对噪声非常敏感,可能会出现特征点误匹配的现象,针对这种情况,在 LK 光流法匹配特征点之后,用 RANSAC 剔除误差较大的特征点,以减小误差^[19]。

1.2 三帧差分法

三帧差分法属于帧间差分法的改进,帧间差分法是一种原理简单,易于实现的运动物体检测方法,该方法运行速度快,受光照影响小。帧间差分法通过从图片序列中选取相邻帧进行差分,比较相邻帧对应位置的像素值,然后设定合适的阈值 T ,将运动物体检测出来^[20]。

假设图像序列中第 K 帧图像为 $f_k(x, y)$, 第 $K-1$ 帧图像为 $f_{k-1}(x, y)$, 那么经过差分后的图像为:

$$D(x, y) = \begin{cases} 1, & |f_k(x, y) - f_{k-1}(x, y)| \geq T \\ 0, & |f_k(x, y) - f_{k-1}(x, y)| < T \end{cases} \quad (1)$$

将差分后的图像与设定的阈值 T 进行比较,超过阈值的像素点归类为前景,否则为背景。三帧差分法与传统帧差法类似,但是能够消除帧间差分法中的“重影”现象。假设连续的三帧图像分别为 $f_{k-1}(x, y)$, $f_k(x, y)$, $f_{k+1}(x, y)$, 对连续的三帧图像两两进行差分运算,对差

2 目标跟踪算法设计

KCF 最早在 2014 年由 Henriques J F 等提出来,该算法使用目标周围区域的循环矩阵采集正负样本,利用岭回归训练目标检测器,并成功地利用循环矩阵在傅里叶空间可对角化的性质将矩阵的运算转化为向量的哈达玛积,同时将线性空间的脊回归通过核函数映射到非线性空间,大大降低了运算量,提高了运算速度,使算法满足实时性要求^[21]。

传统 KCF 算法在跟踪过程当中的目标框是已经设定好的,从始至终大小不发生变化,因此当视频序列中目标尺寸缩小或者发生遮挡时,滤波器就会学习到大量背景信息;如果目标尺寸变大,滤波器只能检测到目标的局部纹理,这几种情况都很可能出现非预期的结果,导致目标框漂移。

本文在 KCF 算法的基础上进行了改进,在进行目标检测时,引入尺度池技术和遮挡检测算法,以改善算法的跟踪效果。

设训练样本集为 (x_i, y_i) , 要找到一个函数 $f(z) = \omega^T z$, 使样本 x_i 和回归目标 y_i 之间的均方误差最小化,其中 ω 是列向量的权重系数,即:

$$\min_{\omega} \|\mathbf{x}\omega - \mathbf{y}\|^2 + \lambda \|\omega\|^2 \quad (2)$$

式中: λ 为惩罚项的系数, 令式(2)导数为 0, 可得:

$$\omega = (\mathbf{x}^H \mathbf{x} + \lambda \mathbf{I})^{-1} \mathbf{x}^H \mathbf{y} \quad (3)$$

式中: $\mathbf{x}^H$ 表示复共轭转置矩阵; $\mathbf{I}$ 为单位矩阵。为了对式(3)进行快速求解, 引入了循环矩阵。KCF 中所有的训练样本是由目标样本循环移位得到的, 假设有一个 $1 \times n$ 的列向量, 将该向量循环移位, 就可以得到一个 $n \times n$ 的循环矩阵, 矩阵定义如式(4)所示。

$$\mathbf{C}(\mathbf{x}) = \begin{bmatrix} x_0 & x_1 & \cdots & x_{n-1} \\ x_{n-1} & x_0 & \cdots & x_{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ x_1 & x_2 & \cdots & x_0 \end{bmatrix} \quad (4)$$

可以证明, 所有的循环矩阵都能够傅氏空间中使用离散傅里叶矩阵进行对角化, 即:

$$\mathbf{x} = \mathbf{F} \text{diag}(\hat{\mathbf{x}}) \mathbf{F}^H \quad (5)$$

式中: $\hat{\mathbf{x}}$ 为 $\mathbf{x}$ 的傅里叶变换, $\hat{\mathbf{x}} = f(\mathbf{x}) = \sqrt{n} \mathbf{F} \mathbf{x}$, $\mathbf{F}$ 是离散傅里叶矩阵常量。

$$\mathbf{F} = \frac{1}{\sqrt{n}} \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ 1 & \omega & \cdots & \omega^{n-2} & \omega^{n-1} \\ 1 & \omega^2 & \cdots & \omega^{2(n-2)} & \omega^{2(n-1)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & \omega^{n-1} & \cdots & \omega^{(n-1)(n-2)} & \omega^{(n-1)(n-1)} \end{bmatrix} \quad (6)$$

将式(6)代入岭回归公式可得:

$$\omega = \mathbf{F} \text{diag} \left(\frac{\hat{\mathbf{x}}^*}{\hat{\mathbf{x}}^* \circ \hat{\mathbf{x}} + \lambda} \right) \mathbf{F}^H \mathbf{y} \quad (7)$$

式中: $\odot$ 表示矩阵的哈达玛积运算。由于:

$$f(\mathbf{C}(\mathbf{x}) \mathbf{y}) = f^*(\mathbf{x}) \circ f(\mathbf{y}) \quad (8)$$

对上式两边同时傅氏变换得:

$$\hat{\omega} = f(\omega) = f^* \left(f^{-1} \left(\frac{\hat{\mathbf{x}}^*}{\hat{\mathbf{x}}^* \circ \hat{\mathbf{x}} + \lambda} \right) \right) \circ f(\mathbf{y}) \quad (9)$$

进一步可得:

$$\hat{\omega} = \frac{\hat{\mathbf{x}}^* \hat{\mathbf{y}}}{\hat{\mathbf{x}}^* \circ \hat{\mathbf{x}} + \lambda} \quad (10)$$

这样就可以使用向量的点积运算取代矩阵运算, 大大提高了计算速度。上面推导的是线性空间的岭回归, 下面将线性空间的岭回归通过核函数映射到非线性空间。设输入图像块为 z , 分类器响应为:

$$f(z) = \omega^T \phi(z) \quad (11)$$

式中: $\phi(\cdot)$ 为非线性映射函数。

$$\omega = \sum_i \alpha_i \phi(x_i) \quad (12)$$

则原线性回归问题可以表示为:

$$f(z) = \sum_{i=1}^n \alpha_i \kappa(z, x_i) \quad (13)$$

式中: $\kappa(z, x_i) = \phi^T(z) \phi(x_i)$ 称为核函数。同样地, 可以求出:

$$\hat{\alpha} = \frac{\hat{\mathbf{y}}}{\hat{\kappa}^{zz} + \mathbf{I} \lambda} \quad (14)$$

式中: κ^{xz} 为核矩阵 κ 的第 1 行, $\kappa = \phi(x) \phi(z)^T$; α 为线性组合系数 α_i 组成的向量。取候选目标的特征为 z , 维度和 x 相同, 推导得出:

$$\hat{f}(z) = \hat{\kappa}^{xz} \circ \hat{\alpha} \quad (15)$$

设目标区域大小为 $s_T = (s_x, s_y)$, 令尺度池为 $\mathbf{s} = [t_1 \ t_2 \ \cdots \ t_n]$, 在当前帧中, 将尺度池作用于目标区域, 令:

$$z_{t_i} = t_i \cdot z (t_i \in \mathbf{s}) \quad (16)$$

依次计算滤波器在 n 个不同尺寸的目标区域中的响应值, 并选择响应最大的尺度对候选区域进行更新, 即:

$$\text{ROI} = \max[f^{-1}(\hat{f}(z_{t_i}))] \quad (17)$$

由于 KCF 算法采用线性插值的方法进行模型更新, 当目标发生部分遮挡时, 会将错误的目标信息加入到模型中, 使得跟踪误差积累, 最终导致跟丢目标, 其计算公式如下:

$$\alpha_t = (1 - \eta) \alpha_{t-1} + \eta \alpha_t, x_t = (1 - \eta) x_{t-1} + \eta x_t \quad (18)$$

式中: η 是插值系数; α 是滤波器系数; x 指目标模板; $t-1$ 代表前一帧的状态; t 代表当前帧的状态。

为了减少部分遮挡和短时完全遮挡对跟踪算法的影响, 本文在算法中加入了遮挡检测机制, 当检测到发生遮挡时, 停止模型更新。如图 3 所示, 该视频序列为 OTB-Jogging, 图 3(a) 所示为目标未受遮挡时的原图与响应图, 图 3(b) 所示为目标受到遮挡时的原图与响应图, 矩形区域中为跟踪的目标, 在响应图中, 颜色越亮代表模型在该点的响应越大。根据响应图, 可以很容易判断目标是否发生遮挡。遮挡检测算法如下: 1) 求响应图中的最大值 $f_{\max}(z)$ 和最大值的位置 $p_{\max}(z)$; 2) 求响应图中像素值大于 $\lambda_1 f_{\max}(z)$ 的所有位置 $\text{pos}(z)$; 3) 以 $p_{\max}(z)$ 为中心, $\lambda_2 \cdot H$ 为高, $\lambda_3 \cdot W$ 为宽构建矩形 R , 判断 $\text{pos}(z)$ 中落入 R 中的比例, 当比例小于 0.8 时, 则判断为发生遮挡, 反之则未发生遮挡。其中 $\lambda_1, \lambda_2, \lambda_3$ 为比例因子, 本文中取 $\lambda_1 = 0.35, \lambda_2 = 0.125, \lambda_3 = 0.24$ 。

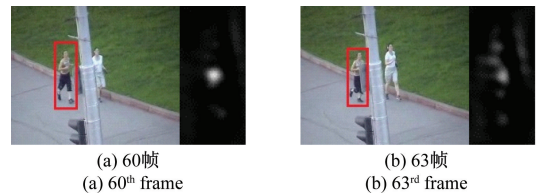


图 3 遮挡与未遮挡条件下的模型响应

Fig.3 The model responses under occluded and unoccluded conditions

改进 KCF 的算法流程如图 4 所示, 首先框选出候选

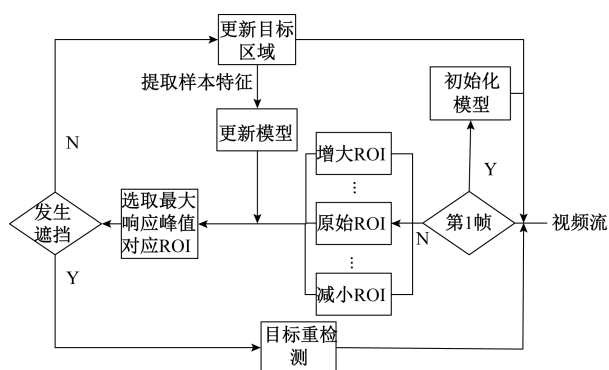


图4 改进KCF的算法流程

Fig.4 The flow chart of the improved KCF algorithm

区域(region of interest, ROI), 然后对 ROI 进行特征提取, 使用目标周围区域的循环矩阵产生正负样本, 采用脊回归训练目标检测器, 实时更新模型。

当利用训练好的检测器对目标进行检测时, 设第 i 帧时目标尺寸为 $s_{Ti} = (s_x, s_y)$, 此时滤波器在候选区域达到最大响应值, 第 $i+1$ 帧时, 对候选区域进行尺度变化, 获得目标的金字塔图像, 记为 $f_1, f_2, \dots, f_n$, 依次计算滤波器在 $f_1, f_2, \dots, f_n$ 中的响应值, 记为 $g_1, g_2, \dots, g_n$, 然后比较各个响应值的大小, 选择响应最大的尺度对候选区域进行更新, 以解决目标跟踪过程当中目标尺度变化的问题。

接着根据滤波器响应图中的灰度分布空间特征判断目标是否被遮挡, 如果发生遮挡, 则停止模型的更新, 继续使用遮挡前的目标模板和滤波器系数, 并且加入目标重检测机制, 为了加快算法的运行速度, 在目标遮挡时记录当前帧中的目标中心位置 (x, y) 以及目标框的宽 w 和高 h , 重检测时, 以 (x, y) 为中心且宽和高分别为 $\lambda w, \lambda h$ 的矩形作为搜索区域, 寻找目标的位置, 没找到目标时, 不对目标框进行更新。

3 仿真结果与分析

3.1 稳像效果分析

本文测试平台配置如下: Intel(R) Core(TM) i7-7700@3.60 GHz, 16 G 内存, Windows7 操作系统。仿真所使用的视频来源为 OTB 数据集^[22], OTB 数据集所用视频都经过人工标注, 分为 11 个属性, 如遮挡、快速移动、光照变化等。

为了验证稳像算法的可靠性, 对抖动干扰下的视频进行了稳像性能测试, 并对稳像前后的效果根据峰值信噪比 (peak signal to noise ratio, PSNR) 进行了比较。PSNR 目前广泛的应用于衡量图像质量, 是衡量图像失真或是噪声水平的客观标准^[23]。两个图像之间 PSNR

值越大, 表明视频连续帧间重合度越好, 视频越稳定。

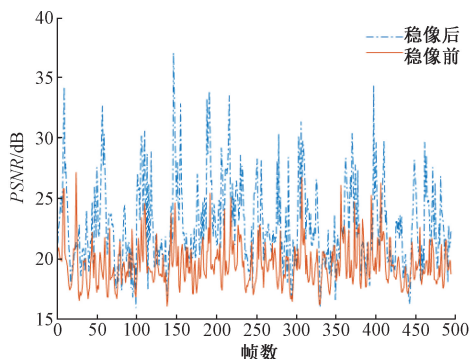
令 $f(x, y)$ 为参考图像, $f'(x, y)$ 为重建图像, 两幅图像 (尺寸为 $M \times N$) 之间的均方根误差为:

$$e_{\text{rms}} = \sqrt{\frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [f'(x, y) - f(x, y)]^2} \quad (19)$$

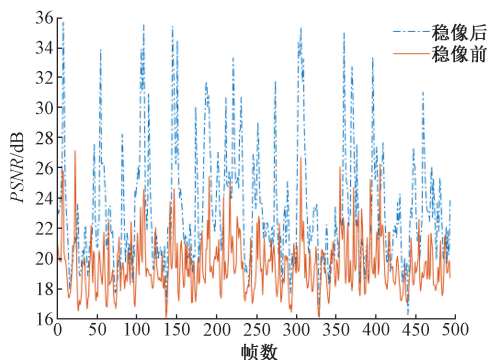
则 PSNR 定义为:

$$\text{PSNR} = 10 \log \frac{(L-1)^2}{(e_{\text{rms}})^2} \quad (20)$$

式中: L 为灰度级数。图 5 所示为稳像前后图像序列的 PSNR 值, 视频源为 OTB-BlurCar1, 并选取了前 500 帧进行测试, 其中纵坐标是视频序列的 PSNR 值, 横坐标是视频的帧数。由图 5 可知, 稳像后的 PSNR 值有了明显的提高。经过计算, 采用 SURF 特征值作为特征点, PSNR 均值提高了 3.2 dB; 采用 SURF 特征点, 并使用 RANSAC 算法剔除部分误匹配点, PSNR 均值提高了 3.6 dB。



(a) SURF算法稳像结果
(a) The image stabilization result for the SURF algorithm



(b) SURF+RANSAC算法稳像
(b) The image stabilization result for the SURF+RANSAC algorithm

图5 各稳像算法 PSNR 曲线

Fig.5 The PSNR curves of different image stabilization algorithms

图 6 所示为实地拍摄获取的抖动视频进行稳像处理的情况, 视频总帧数为 860 帧, 图像尺寸为 640×480 像素。稳像效果如图 6 所示, 图 6(a) ~ (d) 所示分别为未经稳像处理的视频序列, 图 6(e) ~ (h) 所示分别为经过稳像算法处理的图像序列, 可以明显看出, 稳像前目标在

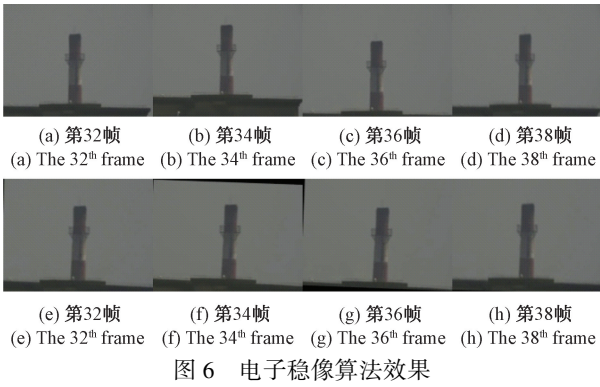


Fig.6 The effects of the electronic image stabilization algorithm

上下抖动,且抖动幅度较大,稳像后目标抖动幅度较小。

3.2 目标检测效果分析

本部分对稳像前后的检测结果进行比较,并将三帧差分算法检测结果与传统的帧间差分法进行了比较,目标检测结果如图 7 所示。在图 7(a)中,未对视频进行稳像,抖动干扰对前景提取影响很大,由于抖动的影响,背景被错误的当作目标提取出来,虚警率较高;图(b)采用帧间差分法,并在处理前对视频做了稳像处理。由结果可以看出,图 7(b)的检测结果较图 7(a)有了明显提高,但是图中出现了“重影”问题;图 7(c)采用了三帧差分法,比较有效地解决了重影问题,但是检测的目标不完整,内部含有“空洞”;图 7(d)对三帧差分后的图像进行了膨胀操作,并寻找膨胀后图像的连通域,求出最小外接矩形。

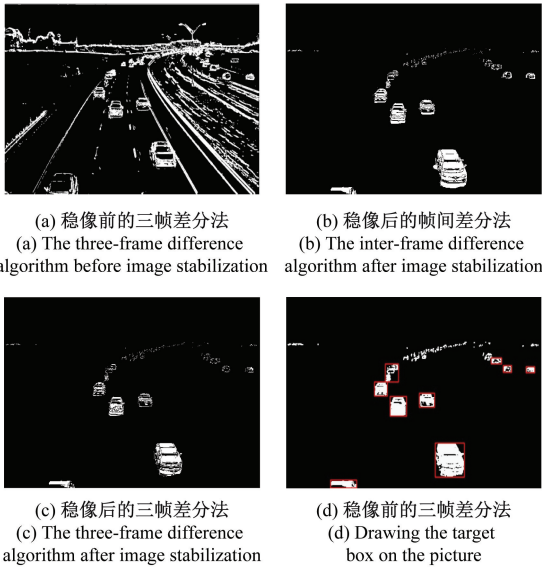


图 7 三帧差分法目标检测结果

Fig.7 The target detection result of three-frame difference algorithm

为了测试基于稳像的目标检测算法在测试平台上的平均处理速度,将算法在 4 组视频上进行了测试,每个视

频选取 400 帧,视频每帧的大小为 640×480 像素,结果显示算法在测试平台上的平均处理速度为 28 fps。

3.3 目标跟踪效果分析

为了验证所设计算法对目标跟踪的准确性和实时性,选取了 OTB100 数据集上的视频对算法进行了测试,并与其他算法进行了比较分析。

1) 像素误差

为了获取不同算法的跟踪像素误差,本文采取了如下操作:(1) 获取原视频每一帧的矩形目标框对角线两点(矩形左上角和右下角)的坐标,记为 $R_1 = (x_1, y_1)$, $R_2 = (x_2, y_2)$; (2) 使用传统 KCF 算法与本文算法,对目标进行跟踪,并获取每帧的目标框的对角线坐标,记为 $R_{KCF1}, R_{KCF2}, R_{ours1}, R_{ours2}$; (3) 计算对应点之间的距离并作图。定义 $\|X_1, X_2\|$ 为两点间的欧氏距离,得到本文算法的像素误差为:

$$D_{ours} = \|R_{ours1}, R_1\| + \|R_{ours2}, R_2\| \quad (21)$$

同理,可得 KCF 算法像素误差 D_{KCF} ,以 D 为纵坐标(单位为像素),帧数为横坐标作图,如图 8 所示,可以看出本文的跟踪算法较传统的 KCF 算法,准确度明显提高。

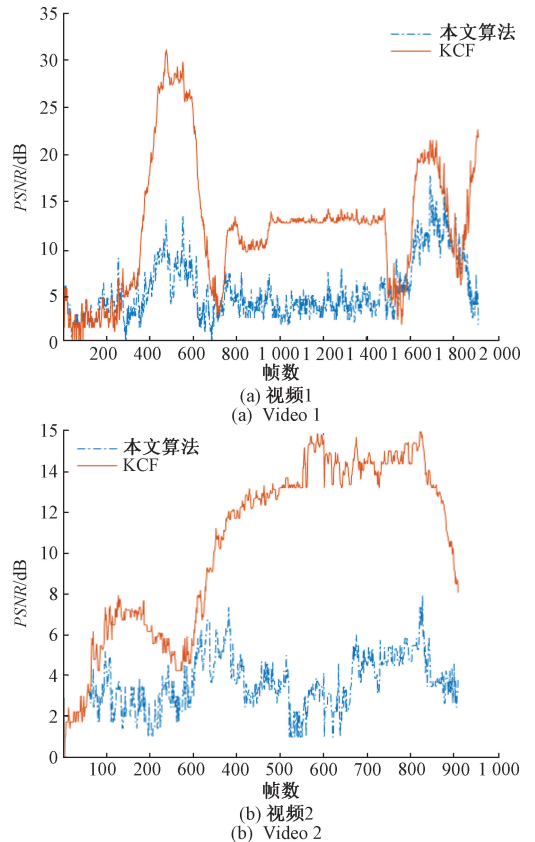


图 8 传统 KCF 算法与改进的 KCF 算法比较

Fig.8 Comparison between traditional KCF algorithm and the improved KCF algorithm

以视频1为例,视频1是 OTB100-RedTeam,摄像头与目标之间距离反复变化。400 和 1 600 帧左右时,由于摄像头由远及近靠近目标,目标慢慢变大,传统 KCF 算法检测误差慢慢变大,误差峰值为 30 pixel;在 1 900 帧左右时,目标由近及远远离摄像头,传统 KCF 算法误差也明显提高。由此可见,传统 KCF 算法对于尺度变化的目标跟踪性能较差,而本文设计的改进算法,在整个视频序列中的误差均在 15 pixel 以下。

视频2为 OTB100-Car2,目标在视频中存在部分遮挡,本文所设计的算法误差稳定在 8 像素以下,传统 KCF 算法在 500~800 帧之间误差较大,达到了 16 pixel。可见本文所设计的跟踪算法对尺度变化和部分遮挡目标的跟踪效果良好。

2) 一次性评估 (one pass evaluations, OPE) 仿真

(1) 准确率

准确率定义为跟踪算法估计的矩形区域的中心点与人工标注区域的中心点之间的距离小于 20 pixel 的视频帧所占百分比。

(2) 成功率

成功率定义为跟踪算法估计的矩形区域与人工标注区域之间的重合率大于 0.5 的视频帧所占百分比。记重合率为 OS ,跟踪算法估计的矩形区域为 A ,人工标注区域为 B ,则:

$$OS = \frac{A \cap B}{A \cup B}$$

(22)

结果表明,本文算法在测试平台上的平均处理速度为 33 fps,在 OTB100 数据集上的 OPE 仿真结果如图 9 和表 1 所示。准确率为 78.2%,相较于 KCF 提高了 4.2%,成功率为 74.1%,相较于 KCF 提高了 11.8%。

表 1 各跟踪算法比较

Table 1 Comparison of various tracking algorithms

OPE 仿真模型	准确率	成功率
SRDCF	0.870	0.814
LCT	0.848	0.813
ECO-HC	0.866	0.810
LMCF	0.842	0.800
SRDCF	0.838	0.781
Staple	0.793	0.754
ours	0.782	0.741
SAMF	0.785	0.732
DSST	0.739	0.670
KCF	0.740	0.623

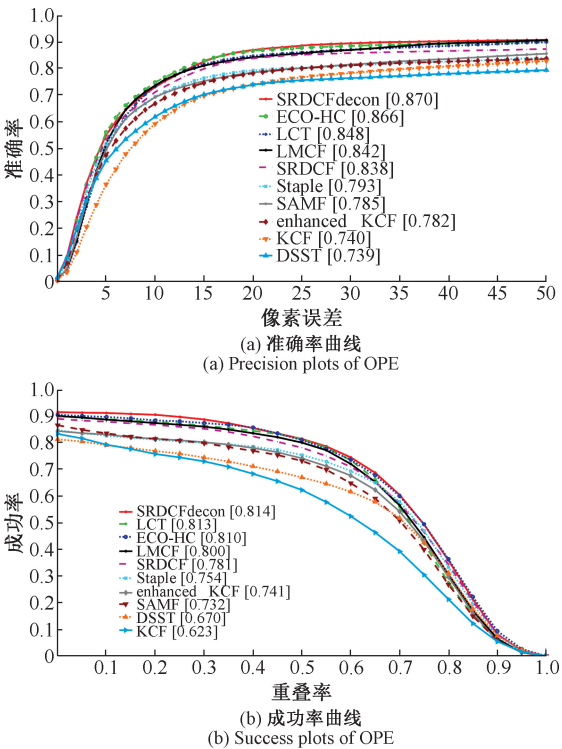


图 9 改进 KCF 算法在 OTB100 数据集上的 OPE 结果
Fig.9 The OPE result of the improved KCF algorithm on OTB100 dataset

4 结 论

本文针对复杂背景下抖动视频序列的目标检测问题,设计了一种结合 LK 光流法与三帧差分法的目标检测算法。该方法能够有效地对抖动视频进行去抖,稳像后的 PSNR 均值提高了 3.6 dB 左右,在稳像的基础上,运用三帧差分法准确的提取出了前景目标,对于像素大小为 640×480 的图片,算法在测试平台上的平均处理速度为 28 fps;同时针对传统 KCF 算法对尺度变化和部分遮挡目标跟踪效果不好的问题,引入了尺度池和遮挡检测算法,设计了一种改进的 KCF 算法。仿真结果表明,本文所设计的跟踪算法,对尺度变化和部分遮挡目标的跟踪效果良好,算法在测试平台上的平均处理速度为 33 fps。

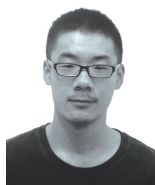
参考文献

[1] 孙宇嘉,于纪言,王晓鸣. 适用于复杂场景的多目标跟踪算法 [J]. 仪器仪表学报, 2019, 40(3) :126-137.
SUN Y J, YU J Y, WANG X M. Multiple object tracking algorithm for the complex scenario [J]. Chinese Journal of Scientific Instrument, 2019, 40(3) :126-137.

- [2] 汪济洲, 鲁昌华, 蒋薇薇. 一种基于随机场多运动目标跟踪算法 [J]. 电子测量与仪器学报, 2017, 31(6): 909-913.
- WANG J ZH, LU CH H, JIANG W W. New multi-target tracking algorithm based on conditional random field [J]. Journal of Electronic Measurement and Instrumentation, 2017, 31(6): 909-913.
- [3] WAN M J, GU G H, CAO E C, et al. In-frame and inter-frame information based infrared moving small target detection under complex cloud backgrounds [J]. Infrared Physics and Technology, 2016, 76(5): 455-467.
- [4] 华媛蕾, 刘万军. 改进混合高斯模型的运动目标检测算法 [J]. 计算机应用, 2014, 34(2): 580-584.
- HUA Y L, LIU W J. Moving object detection algorithm of improved gaussian mixture model [J]. Journal of Computer Applications, 2014, 34(2): 580-584.
- [5] GU B, SONG K F, QIU D Y, et al. Moving object detection based on improved ViBe algorithm [J]. International Journal of Smart Home, 2015, 9(12): 225-232.
- [6] GAO P, SUN X G, WANG W. Moving object detection based on kirsch operator combined with optical flow [C]. international Conference on Image Analysis and Signal Processing, 2010: 620-624.
- [7] 尹宏鹏, 陈波, 柴毅, 等. 基于视觉的目标检测与跟踪综述 [J]. 自动化学报, 2016, 42(10): 1466-1489.
- YIN H P, CHEN B, CHAI Y, et al. Vision-based object detection and tracking: a review [J]. Acta Automatica Sinica, 2016, 42(10): 1466-1489.
- [8] 葛宝义, 左宪章, 胡永江. 视觉目标跟踪方法研究综述 [J]. 中国图象图形学报, 2018, 23(8): 1091-1107.
- GE B Y, ZUO X ZH, HU Y J. Review of visual object tracking technology [J]. Journal of Image and Graphics, 2018, 23(8): 1091-1107.
- [9] TUZEL O, PORIKLI F, MEER P. A bayesian approach to background modeling [C]. Computer Society Conference on Computer Vision and Pattern Recognition, 2005: 58-65.
- [10] 吴孟俊, 付钊, 刘建平, 等. 基于动态贝叶斯网络的多特征目标跟踪 [J]. 计算机工程与应用, 2011, 47(30): 183-187.
- WU M J, FU T, LIU J P, et al. Multi-feature target tracking based on dynamic bayesian network [J]. Computer Engineering and Applications, 2011, 47(30): 183-187.
- [11] 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, 2015, 37(3): 583-596.
- [12] KALAL Z, MIKOLAJCZYK K, MATAS J. Tracking-learning-detection [J]. IEEE Transactions on Software Engineering, 2011, 34(7): 1409-1422.
- [13] HARE S, TORR P H S. Struck: Structured output tracking with kernels [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38(10): 2096-2109.
- [14] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection [C]. Computer Society Conference on Computer Vision and Pattern Recognition, 2005: 886-893.
- [15] LI Y, ZHU J K. A scale adaptive kernel correlation filter tracker with feature integration [C]. Proceedings of the 13th European Conference on Computer Vision, 2014: 254-265.
- [16] ZHU G B, WANG J Q, WU Y, et al. MC-HOG correlation tracking with saliency proposal [C]. Thirtieth AAAI Conference on Artificial Intelligence, 2016: 3690-3696.
- [17] 王志民, 徐晓刚. 电子稳像技术综述 [J]. 中国图象图形学报, 2010, 15(3): 470-480.
- WANG ZH M, XU X G. A survey on electronic image stabilization [J]. Journal of Image and Graphics, 2010, 15(3): 470-480.
- [18] 朱奇光, 张朋珍, 李昊立, 等. 基于全局和局部特征融合的图像匹配算法研究 [J]. 仪器仪表学报, 2016, 37(1): 170-176.
- ZHU Q G, ZHANG P ZH, LI H L, et al. Investigation on the image matching algorithm based on global and local feature fusion [J]. Chinese Journal of Scientific Instrument, 2016, 37(1): 170-176.
- [19] 陈月, 赵岩, 王世刚. 图像局部特征自适应的快速 SIFT 图像拼接方法 [J]. 中国光学, 2016, 9(4): 415-422.
- CHEN Y, ZHAO Y, WANG SH G. Fast image stitching method based on SIFT with adaptive local

- image feature [J]. Chinese Optics, 2016, 9(4): 415-422.
- [20] ZHANG L P, SHAO Z K, WU J D, et al. A moving object detection method based on improved three-frame difference algorithm and edge information [C]. International Conference on Advances in Mechanical Engineering and Industrial Informatics, 2015:279-284.
- [21] 张微, 康宝生. 相关滤波目标跟踪进展综述 [J]. 中国图象图形学报, 2017, 22(8):1017-1033.
- ZHANG W, KANG B SH. Recent advances in correlation filter-based object tracking: a review [J]. Journal of Image and Graphics, 2017, 22(8):1017-1033.
- [22] WU Y, LIM J W, YANG M H. Object tracking benchmark [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1834-1848.
- [23] 万国挺, 王俊平, 李锦, 等. 图像拼接质量评价方法 [J]. 通信学报, 2013, 34(8):76-81.
- WANG G T, WANG J P, LI J, et al. Method for quality assessment of image mosaic [J]. Journal on Communications, 2013, 34(8): 76-81.

作者简介



郑浦, 2017 年于中北大学获得学士学位, 现为南京理工大学硕士研究生, 主要研究方向为图像处理。

E-mail: 117108022106@njust.edu.cn

Zheng Pu received his B.Sc. degree from North University of China in 2017. Now, he is a M. Sc. candidate in Nanjing University of Science and Technology. His main research interest is image processing.



白宏阳 (通信作者), 2012 年于南京理工大学获得博士学位, 现为南京理工大学副教授, 主要研究方向为弹箭组合导航与末制导技术、目标精确识别技术。

E-mail: hongyang@njust.edu.cn

Bai Hongyang (Corresponding author) received his Ph. D. degree from Nanjing University of Science and Technology in 2012. Now, he is an associate professor in Nanjing University of Science and Technology. His main research interests include integrated navigation, terminal guidance technology and precise target recognition for missiles and rockets.