

# 红外图像中快速运动目标的检测与跟踪方法

易诗，张洋溢，聂焱，赵茜茜，庄依彤

(成都理工大学 信息科学与技术学院, 四川 成都 610059)

**摘要:** 红外热成像图像具有分辨率较低, 细节模糊, 对于快速运动目标适应性较差的特点。本文提出了一种结合目标检测算法, 目标跟踪算法的红外图像中快速运动目标的检测与跟踪方法。该方法根据红外图像特点, 使用 ViBE 算法检测运动目标, 检测出图像中显著运动目标后, 触发跟踪器, 使用 fDSST 目标跟踪算法对显著运动目标进行跟踪。测试结果表明, 该方法对于快速运动的红外图像目标能够高效检测、快速跟踪。检测与跟踪效果相对传统方法具有检测率更高、鲁棒性更好、实时性更强的优势, 对于红外图像中目标检测与跟踪具有很强应用价值。

**关键词:** 红外图像; 运动目标检测; 目标跟踪; ViBE 算法, fDSST 算法

中图分类号: TN919.5 文献标识码: A 文章编号: 1001-8891(2019)03-0268-05

## Fast-moving Target Detection and Tracking Method in Infrared Image

YI Shi, ZHANG Yangyi, NIE Yan, ZHAO Qianqian, ZHUANG Yitong

(College of Information Science and Technology, Chengdu University of Technology, Chengdu 610059, China)

**Abstract:** An infrared thermal image exhibits characteristics of low resolution, fuzzy details, and poor adaptability to fast-moving targets. In this paper, a fast-moving target detection and tracking method for an infrared image is proposed, which combines infrared thermal image with machine vision. According to the characteristics of an infrared image, the ViBE algorithm is used to detect moving objects. After detecting the salient moving objects in the image, the tracker is triggered and the fast discriminative scale space tracking algorithm for target tracking is used to track the salient moving objects. The test results demonstrate that this method can detect and track the fast-moving infrared image target efficiently, and its detection and tracking efficiency is higher than in the traditional method. In addition, it is robust and has real-time advantages. Therefore, this method has a strong application value for infrared image target detection and tracking.

**Key words:** infrared image, moving target detection, target tracking, ViBE algorithm, fDSST algorithm

## 0 引言

当夜间, 目标的颜色、形态与周围环境相似的情况下, 红外热成像成为了搜索目标的重要手段。当夜间, 目标的颜色、形态与周围环境相似的情况下, 红外热成像成为了搜索目标的重要手段。它广泛运用于安防监控、军事侦查、夜间行车、航运等领域。红外图像反应的是物体的相对温度信息。红外光谱成像受天气因素的影响较小。在无光与雨雾天气, 相对于照明摄像头、夜视等设备红外光谱成像具有探测距离远、探测可靠性高等优势<sup>[1]</sup>。但红外光谱成像具有分

辨率较低, 细节模糊等缺点。红外图像中如背景存在多个热源会造成目标与背景差异较小, 对于快速运动目标的适应性较差<sup>[2]</sup>。传统方法对于红外图像目标的检测与跟踪, 大多采用热源跟踪方法。该方法利用锁定热源最明显的区域进行跟踪, 当热源出现较多时, 此类方法存在误跟踪, 鲁棒性较差的缺陷。而且在传统红外图像跟踪方法中缺少运动检测环节, 不能有效地检测出红外图像中的运动目标。

本文设计实现了一种针对红外图像中快速运动目标的检测与跟踪方法, 该方法使用目前效果较好的 ViBE 运动目标检测算法进行运动目标检测<sup>[2]</sup>, 检测出

收稿日期: 2018-09-11; 修订日期: 2018-12-21。

作者简介: 易诗 (1983-), 男, 四川成都人, 硕士研究生, 高级实验师, 主要从事机器视觉研究, 深度学习算法研究, 信号与信息处理研究。E-mail: 549745481@qq.com。

基金项目: 国家大学生创新创业项目 (201810616033)。

图像中显著运动目标后,触发跟踪器,跟踪算法采用目前速度最快的 fDSST 目标跟踪算法对显著运动目标进行高速、鲁棒性良好的跟踪。相对于传统方法,该方法具备目标检测率高、鲁棒性好、速度快的优势,具有较高的科研和工程应用价值。

## 1 目标的运动检测

该方法首先需检测红外图像中的运动目标,因此,当存在多个热源时,运动目标与背景差异较小,对于红外图像中运动目标的检测需采用运动目标前景提取算法进行高效的检测<sup>[3]</sup>。

本文采用 ViBE 算法进行运动目标检测,该算法为 2011 年提出的一种基于背景建模的运动检测算法<sup>[4]</sup>。

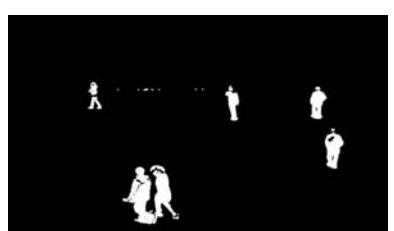
ViBE 算法相对于传统算法具备前景检测和背景模型更优的优势。传统运动目标检测算法中, GMM<sup>[5]</sup> 算法在检测效果和速度方面具有很强代表性,运用广泛。图 1 为 ViBE 算法与 GMM 算法对运动目标检测的比较。



(a) 原始图像 (a) Original image



(b) GMM 算法前景提取  
(b) Foreground extraction of GMM algorithm



(c) ViBE 算法前景提取  
(c) Foreground extraction of ViBE algorithm

图 1 ViBE 算法与 GMM 算法前景提取对比

Fig.1 Comparison of foreground extraction between ViBE algorithm and GMM algorithm

由测试效果可见 ViBE 算法提取运动目标前景效果优于 GMM 运动目标检测算法。

ViBE 算法提取运动目标前景分为 3 个步骤:

1) 初始化单帧图像中每个像素点的背景模型。当输入第一帧图像时,建立像素的背景模型。如式(1)所示:

$$BK_M^0 = f^0(x^i, y^i) | (x^i, y^i) \in N_G(x, y) \quad (1)$$

式中:  $N_G(x, y)$  表示空域上相邻的像素值;  $f^0(x^i, y^i)$  表示当前点的像素值。在初始化过程中,选择第一帧图像建立背景模型。

2) 对后续的图像序列进行前景目标分割操作。

初始化之后,前景提取由下式决定:

$$f^k(x, y) = \begin{cases} BK_M^{k-1} = |(x^r, y^r) > T, & \text{foreground} \\ BK_M^{k-1} = |(x^r, y^r) \leq T, & \text{background} \end{cases} \quad (2)$$

式中: 像素点  $(x, y)$ ; 背景模型为  $BK_M^{k-1}$ ; 像素值为  $f^k(x, y)$ ;  $T$  为预设阈值,根据当前像素点背景模型和阈值的比较确定其为前景或背景。

3) 背景模型更新。通常背景变化不大,所以每次背景模型更新的个数相近。因此将第一帧背景更新的次数作为比较值,符合下面公式则对背景模型进行重新初始化。

$$|\text{Update} - \text{Init}| > 0.3 \times \text{Init} \quad (3)$$

式中: Update 为背景更新个数; Init 为一帧图像更新次数,以此避免由于大面积的光照变化导致的误判。

## 2 提取红外图像中的显著运动目标

使用 ViBE 算法检测到红外图像中运动目标后,由于红外图像中可能存在多个热源干扰,因此本文方法提取最显著的运动目标进行跟踪。

提取显著运动目标流程如图 2 所示。



图 2 提取显著运动目标

Fig.2 Extract significant moving targets

在提取红外图像运动前景后,检测前景中所有连通域,将最大连通域作为图像中最为显著的运动目标。

检测出显著运动目标后以该目标为跟踪目标,触发跟踪器对该目标进行跟踪。

## 3 红外目标跟踪算法

检测到红外图像中显著运动目标后,触发跟踪

器, 使用目标跟踪算法对快速运动目标进行跟踪。

考虑到目标的快速移动, 本文需采用一种高速、鲁棒性强的目标跟踪算法。基于相关滤波 (correlation filter) 的跟踪方法因为速度快、效果好吸引了众多研究者的眼光。相关滤波器在运算中巧妙应用快速傅里叶变换获得了大幅度速度提升, 目前基于相关滤波的拓展方法也有很多, 包括核化相关滤波器 KCF, 加尺度估计的相关滤波器 (DSST) 等。近期出现的基于 DCF 对 DSST 改进的 fDSST, 相比于 KCF 和 DSST, 计算速度更快、效率更高、精度更高、鲁棒性更强<sup>[6]</sup>。

表1为国际机器视觉大会上 fDSST 算法与目前流行的跟踪器在 OTB 数据集中 50 个视频进行测试的平均重叠精度, 距离精度, 以及算法运行测试速度 FPS 的比较。

表1 fDSST 算法与流行目标跟踪算法性能比较

Table 1 Performance comparison between fDSST algorithm and popular target tracking algorithm

	CT	TLD	STRUCK	KCF	fDSST
Average					
overlapping accuracy	24.8%	48.9%	58.3%	64.3%	74.3%
Distance accuracy					
Test speed	32.4%	55.3%	68.7%	74%	80.2%
Test speed	67.5	23.6	10.4	104	174

由该数据可见, fDSST 算法在平均重叠精度以及距离精度上优于其他流行目标跟踪算法, 在平均处理速度上更是高达 174 fps。尤其在尺度变化视频中, fDSST 算法明显优于所有跟踪器。因此适宜于跟踪红外图像中快速运动的目标。

fDSST 目标跟踪算法是由加入尺度变化和特征降维的 moose 相关滤波器改进而来<sup>[7]</sup>。

moose (Minimum Output Sum of Squared Error Filter) 相关滤波器工作原理为对  $t$  个目标区域样本分别为  $f_1, f_2, \dots, f_t$  通过滤波器  $h_t$ , 期望输出  $g_i$  (通常为二维高斯函数, 峰值位于目标区域中心), 最小化均方误差:

$$\varepsilon = \sum_{j=1}^t \|h_t f_j - g_j\|^2 = \frac{1}{MN} \sum_{j=1}^t \|\bar{H}_t F_j - G_j\|^2 \quad (4)$$

对  $H_t$  求导, 令导数为 0, 得:

$$H_t = \frac{\sum_{j=1}^t \bar{G}_j F_j}{\sum_{j=1}^t \bar{F}_j F_j} \quad (5)$$

对于  $t+1$  帧, 假设其在第  $t$  帧的目标区域内提取得到特征图  $z$ , 那么计算  $h_t$  与  $z$  的相关值:

$$y = F^{-1}\{\bar{H}, z\} \quad (6)$$

式中:  $y$  的最大值的位置即被认为是  $t+1$  帧中目标区域的中心点。保持目标区域尺寸不变, 将其中心点进行移动相应位置, 就得到了在新帧中的目标区域。对该区域进行特征提取, 然后加入到训练集中对模型进行更新得到  $h_{t+1}$  后, 即可进行下一帧中的目标检测。

moose 相关滤波器跟踪的缺点在于只使用了灰度作为特征, 模型所使用的特征维数太低, 难以很好地反映目标的特性。算法估算了目标区域中心点在帧间的平移运动, 而没有考虑目标在运动过程中反映在画面上的尺度变化, 在目标尺度发生改变时难以适应。

在 moose 相关滤波器的基础上, 发展出了 dsst 目标跟踪算法, 加入了尺度变换, 使用 fhog 替代灰度 (通常 fhog 为 31 dims, 而 dsst 实际使用 28 dims), 考虑目标尺度变化, 提高了跟踪的鲁棒性。

而 fDSST 对 dsst 进行加速, 分别对位置滤波器和尺度滤波器进行 PCA 降维和 QR 分解来降低计算量提高计算速度, 使其更适宜于快速运动目标的跟踪。

fDSST 算法对 DSST 算法的特征降维, 首先对位置滤波器输入  $u_t$  进行降维预测:

$$u_t = (1 - \eta)u_{t-1} + \eta f_t \quad (7)$$

式中:  $f_t$  为当前样本;  $\eta$  为学习率。

预测出位置滤波器输入后, 采集  $n$  帧的数据, 代入相关滤波器求目标峰值, 峰值计算函数中加入了对尺度估算因子  $P_t$ :

$$\varepsilon = \sum_{t=1}^n \|u_t - P_t^T P_t u_t(n)\|^2 \quad (8)$$

在尺度搜索中, fDSST 只考虑 17 个尺度, 而特征有上千, 所以可以将维度压缩到 17 而无信息损失。使用 QR 分解, 基于以上两个加速方法扩大搜索区域。dsst 算法中考虑了 33 个尺度而 fDSST 中只考虑了 17 个尺度, 所以在响应图中插值到 33 个, 使用三角插值<sup>[11]</sup>。

经过一系列改进, fDSST 算法可以高速, 稳定地跟踪快速移动、尺度变化、形状变化的目标。

#### 4 实验结果分析

实验测试采用普雷德 G35 户外热像仪, 其基本参数为: 35 mm 镜头, 17 μm 像元, 探测器分辨率为 384 × 288, 输出 AV 格式视频。采用 USB 单通道 AV 数据采集转换卡读取输出视频并实时转换为 JPG 格式

供电脑处理。视频处理硬件环境为: Intel i7 4 核处理器 2.5 GHz, 8 GB 内存, GeForce GT 950 M 显卡。软件平台使用: VS2015+Opencv3.1+Cuda7.5 环境。

在实际测试的软硬件环境下,首先使用 ViBE 算法进行运动目标检测,选择最大连通域确定显著运动目标。测试结果如图 3 所示。

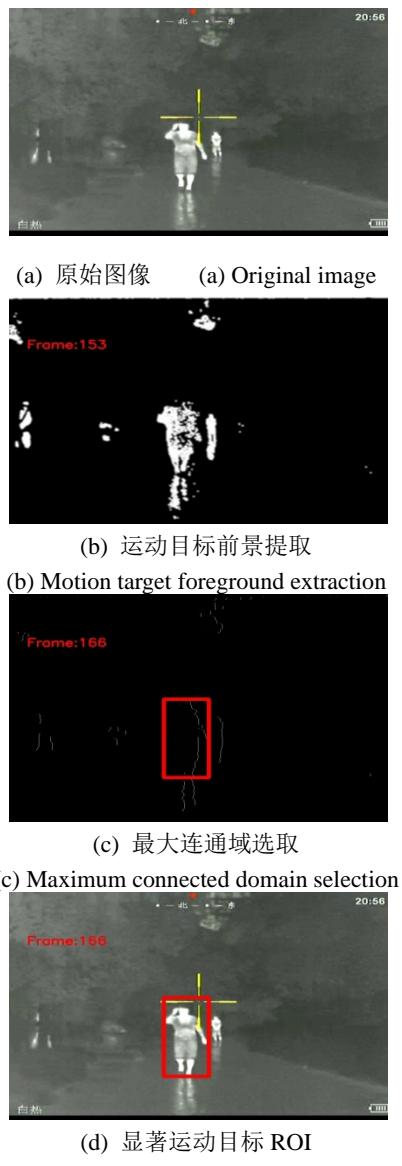


图 3 提取显著运动目标实际测试

Fig.3 Real test for extracting significant moving targets

根据实际测试结果,对显著运动目标的检测时间,提取性能,如表 2 所示。

表 2 提取显著运动目标测试结果

Table 2 Extraction of significant moving target test results

Eliminate ghost	Maximum connected region	Lock significant targets
153f	166f	166f

由测试数据可见,使用 ViBE 算法对显著运动目标提取在第 153 帧时可以有效地消除鬼影,提取清晰的运动目标前景,第 166 帧时检测到最大连通区域并提取到显著运动目标 ROI,有效地检测出了红外图像中显著运动目标。

在选取显著运动目标触发跟踪器后,在跟踪器中使用 fDSST 目标跟踪算法分别与经典目标跟踪算法中的 struck 算法, CT 算法, KCF 算法进行了对比测试,跟踪效果与速度如图 4 所示。

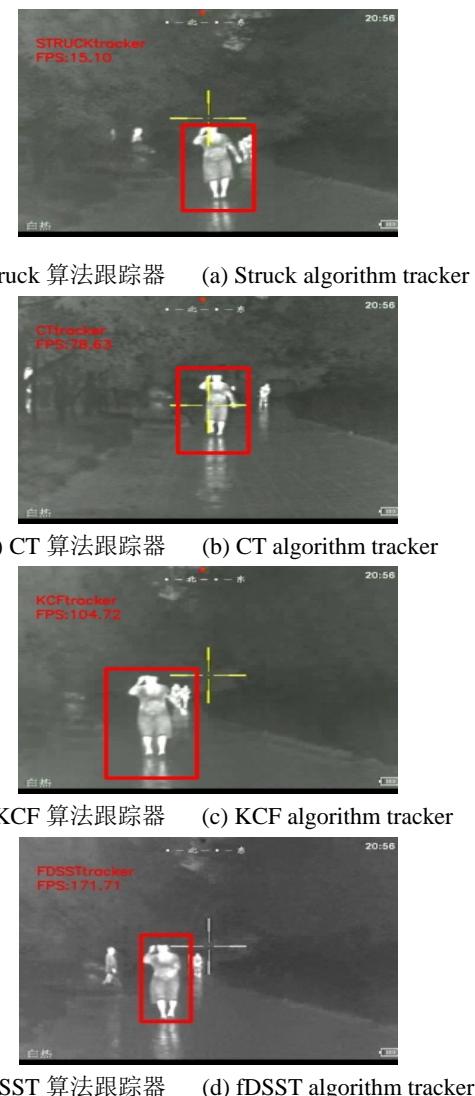


图 4 跟踪算法实际测试对比

Fig.4 Comparison of actual tracking algorithms

对多种跟踪器实际测试数据如表 3 所示。

由实际测试结果可见,相同硬件平台,软件环境下, fDSST 在平均处理速度,目标快速运动,目标尺度变化,目标形状变化等方面相对其余类型采用经典算法的跟踪器具备明显优势。满足高速的同时,鲁棒性更强。适合于红外图像中快速运动目标的跟踪。

表3 跟踪器对比测试 Table 3 Comparison test of trackers

	Average processing speed /fps	Fast Target Target scale change	Target shape change
struck	15.1	Tracking offset	Stable tracking
CT	70.83	Tracking offset	Tracking offset
KCF	104.72	Stable tracking	Tracking offset
fDSST	171.71	Stable tracking	Stable tracking

## 5 结论

本文提出了一种基于 ViBE 运动检测算法与 fDSST 目标跟踪算法的红外图像快速运动目标的检测与跟踪方法。该方法结合了 ViBE 算法良好的运动目标检测性能与 fDSST 高速、稳定的目标跟踪性能，由测试结果看出，该方法能够针对红外图像这类分辨率较低、目标细节模糊的图像中快速运动的目标进行准确、高速地检测与跟踪。

由测试结果看出，该方法仍存在一定的改进空间，可以进一步采用高分辨率、大镜头的热像仪提高原始图像质量，采用更新的图像处理算法提升图像细节。同时，ViBE 算法存在鬼影消除较慢的特点，在今后工作中可以改良这一缺陷，或者采用更新效果更良好的运动目标检测算法。对于目标的跟踪，目前深度学习目标跟踪方法发展迅速，跟踪鲁棒性非常高，但速度较慢，在未来出现基于深度学习框架的高速目标跟踪方法后，可以使用其进行目标跟踪，进一步提高目标跟踪的鲁棒性。

## 参考文献：

- [1] 崔美玉. 论红外热像仪的应用领域及技术特点[J]. 中国安防, 2014(12): 90-93.  
CUI Meiyu. Application field and technical characteristics of infrared thermal imager[J]. *China Security & Protection*, 2014(12): 90-93.
- [2] 范延军. 基于机器视觉的先进辅助驾驶系统关键技术研究[D]. 南京: 东南大学, 2016.  
FAN Yanjun. Research on key technologies of advanced assisted driving system based on machine vision[D]. Nanjing: Southeast University, 2016.
- [3] 张科, 刘彦. 改进的基于背景预测的红外弱小目标检测方法[J]. 火力与指挥控制, 2008, 33(11): 22-24.  
ZHANG Ke, LIU Yan. An Improved Small Target Detection Method Based on background prediction in IR images[J]. *Fire Control and Command control*, 2008, 33(11): 22-24.
- [4] 杨阳, 杨静宇. 基于显著性分割的红外行人检测[J]. 南京理工大学学报: 自然科学版, 2013, 37(2): 251-256.  
YANG Yang, YANG Jingyu. Infrared pedestrian detection based on saliency segmentation[J]. *Journal of Nanjing University of Science and Technology*, 2013, 37(2): 251-256.
- [5] 魏丽, 丁萌, 曾丽君, 等. 红外图像中基于似物性与稀疏编码的行人检测[J]. 红外技术, 2016, 38(9): 752-757.  
WEI Li, DING Meng, CENG Lijun, et al. Pedestrian detection based on quasi physical properties and sparse coding in infrared images[J]. *Infrared Technology*, 2016, 38(9): 752-757.
- [6] Ahmed M N, Yamany S M, Mohamed N, et al. A Modified Fuzzy means Algorithm for Bias Field Estimation and Segmentation of MRI Data[J]. *IEEE Transactions on Medical Imaging*, 2002, 21(3): 193-199.
- [7] ZHENG J, ZHANG D H, HUANG K D, et al. An Adaptive Image Segmentation Method Based on the Fuzzy means with Spatial Information[C]//IET Image Processing, 2017, 12(5): 785-792.
- [8] Henry L, Neville D, NAN X. Detection of small objects in clutter using a GA-RBF neural network[J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2002, 38(1): 98-118.
- [9] Pinngar C R, Mansinha L. Time-local spectral analysis for non-stationary time series: the S-transform for noisy signals[J]. *Fluctuation and noise letters*, 2003, 3(3): 357-364.
- [10] Pinngar C R, Eaton D W. Application of the S transform to prestack noise attenuation filtering[J]. *Journal of Geophysical Research*, 2003, 108(B9): 1-10.
- [11] WU Bo, Nevatia Ram. Detection and tracking of multiple, partially occluded humans by Bayesian combination of edgelet based part detectors[J]. *International Journal of Computer Vision*, 2007, 75(2): 247-266.