

doi: 10.7690/bgzdh.2021.07.009

融合 SVM 的多特征 DSST 目标跟踪算法

王承赞¹, 王思卿², 张龙杰¹, 李彦宽³, 张龙云⁴

(1. 海军航空大学岸防兵学院, 山东 烟台 264001; 2. 中国人民解放军 92555 部队, 上海 201900;
3. 烟台北方星空自控科技有限公司, 山东 烟台 264003; 4. 山东大学后勤保障部, 济南 250000)

摘要: 为解决 DSST 算法多尺度搜索策略跟踪时目标出现严重遮挡、非刚性形变、目标脱离视场导致的目标外观变化的问题, 提出一种将支持向量机 (support vector machine, SVM) 目标重检测模块融合的算法。提取目标的多种特征然后将这些特征矢量融合以增强目标的特征表达。在 DSST 算法的位置和尺度滤波器的基础上, 新增目标外观滤波器, 利用训练好的 SVM 全局搜索目标。采用不同大小的窗口采样来训练相关模型并建立一个 SVM 的最优分类面, 通过 SVM 对丢失后的目标进行重检测。实验结果表明, 改进算法比 DSST 算法在对目标受到遮挡、目标非刚性形变等问题上的鲁棒性能均有提高。

关键词: DSST 算法; 多特征融合; SVM 分类器; 目标重检测; 鲁棒性跟踪

中图分类号: TP391 **文献标志码:** A

Multi-feature DSST Target Tracking Algorithm Based on SVM Fusion

Wang Chengyun¹, Wang Siqing², Zhang Longjie¹, Li Yankuan³, Zhang Longyun⁴

(1. School of Coastal Defense, Navy Aviation University, Yantai 264001, China;

2. No. 92555 Unit of PLA, Shanghai 201900, China;

3. Yantai Beifang Xingkong Self-control Technology Co., Ltd., Yantai 264003, China;

4. Department of Logistics Support, Shandong University, Ji'nan 250000, China)

Abstract: In order to solve the problems of object appearance change caused by severe occlusion, non-rigid deformation and object departure from the field of view in multi-scale search strategy tracking of DSST algorithm, an algorithm was proposed to fuse the support vector machine (SVM) object re-detection module. Multiple features of the target are extracted and these feature vectors are fused to enhance the feature expression of the target. Based on the position filter and scale filter of DSST algorithm, the target appearance filter is added, the trained SVM is used to search for the target globally. Different window samples are used to train relevant models and establish an optimal classification surface of SVM. The missing target is re-detected by SVM classifier. The experimental results show that the improved algorithm has better robust performance than DSST algorithm on such problems as target occlusion and non-rigid deformation.

Keywords: DSST algorithm; multi-feature fusion; SVM classifier; object re-detection; robust tracking

0 引言

目标跟踪是根据已知目标在图像中的位置来预测后续视频序列中目标位置变化, 以达到持续跟踪目标的目的。相关滤波跟踪算法利用样本的循环矩阵对角化的性质, 将计算从时域卷积转换到傅里叶域的矩阵元素点乘, 结合快速傅里叶变换, 极大提升了跟踪速度。而且通过循环移位来采样能够得到大量的训练样本, 增强了相关滤波器对目标的判别能力, 提高了算法的跟踪精度。

相关滤波算法的核心思想是通过采样图像中的目标区域, 计算相邻两帧图像中目标区域的相似度, 与上一帧图像目标区域相似度最大的位置即可认为是新一帧图像的目标区域。MOSSE (minimum output sum of squared error) 算法^[1]是相关滤波的开山之

作, 通过采样目标的外观特征来更新位置滤波器, 预测出目标位置。CSK (circulant structure kernel) 算法^[2]采样时通过循环移位的方法创造出了很多的背景一负样本, 利用目标的正样本和大量的负本来训练模型, 使得相关滤波器从背景更好地判别出目标; KCF (kernelized correlation filters) 算法^[3]利用多通道 (histogram of oriented gradient, HOG) 特征取代原始灰度特征, 并引入高斯核函数, 将线性空间中不可分的样本特征映射入高维非线性空间中, 使样本能够很好的分类; DSST (discriminative scale space tracker) 算法^[4]在位置滤波器的基础上, 又增加了一个尺度滤波器, 通过位置和尺度滤波器分别估计出跟踪框的位移和缩放, 能很好地处理大小变化的目标跟踪问题。

收稿日期: 2021-03-29; 修回日期: 2021-05-02

基金项目: 国家自然科学基金 (51809156); 中国博士后科学基金 (2016M600537)

作者简介: 王承赞 (1992—), 男, 山东人, 硕士, 从事兵器科学与技术研究。E-mail: 2275339722@qq.com。

1 DSST 算法概述

1.1 目标的位置跟踪

DSST 算法用岭回归方法建立目标损失函数:

$$\varepsilon = \left\| \sum_{l=1}^d h^l \otimes f^l - g \right\|^2 + \lambda \sum_{l=1}^d \|h^l\|^2 \quad (1)$$

式中: 考虑到 31 维的 HOG 特征, f^l 表示第 l 维通道的样本特征; h^l 表示第 l 维通道的位置滤波器模型; λ 是正则项, 防止样本过拟合; g 是 2 维的高斯函数, 表示目标的理想输出。

式(1)用到了第 1 维通道的样本特征, 通过最小二乘求解模型在傅里叶域表示:

$$H^l = A^l / B = \bar{G} F^l / \left(\sum_{k=1}^N \bar{F}^k F^k + \lambda \right) \quad (2)$$

式中: N 是训练样本的总数; \bar{G} 表示 2 维高斯标签 G 的复共轭。

对新一帧图像, 用采样窗口提取图像中的测试样本与训练好的模型相关操作, 计算得到其在傅里叶域表示的响应值:

$$\hat{y}(f, z) = \sum_{l=1}^d \bar{A}^l Z^l / (B + \lambda) \quad (3)$$

将响应值通过傅里叶逆变换转到时域, 响应值最大的位置作为目标新位置。

新的一帧图像中, 以目标新位置为中点, 窗口大小是目标跟踪框 2.8 倍的区域上采样, 结合位置模型的历史参数来更新新一帧图像的模型参数:

$$\begin{aligned} A_t^l &= (1-\eta)A_{t-1}^l + \eta\bar{G}F_t^l, \\ B_t &= (1-\eta)B_{t-1} + \eta\bar{F}_t^k F_t^k, \\ H_t^l &= A_t^l / B_t. \end{aligned} \quad (4)$$

式中: η 为更新率; 下标 $t-1$ 为模型的历史参数; 下标 t 为当前帧的模型参数; H_t^l 为当前帧更新后的模型参数。在目标跟踪中目标外观是会发生变化的, DSST 算法对每帧图像提取训练样本, 并结合历史模型, 采用线性加权的方法来更新当前帧图像的模型, 避免模型更新的剧烈程度。

1.2 目标的尺度估计

尺度跟踪的思路与位置跟踪相似, 建立一个尺度池 $S, S = \{s_1, s_2, s_3, \dots, s_s\}$ 。1 维的尺度滤波器以目标的新位置为中心, 在窗口大小分别是跟踪框大小的 a^n 倍的区域上采样 s 个尺度的测试样本集 $\{Z_i | i \in a^n P * a^n R\}$ 。其中: i 表示测试样本的不同尺度; a 是尺度步长; P 和 R 分别是跟踪框的宽和高; n

是尺度的取值范围:

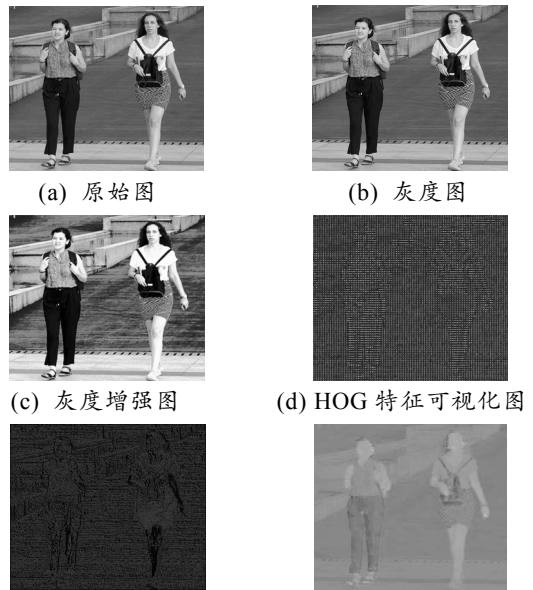
$$n \in \{-(s-1)/2, -((s-3)/2), \dots, (s-1)/2\} \quad (5)$$

尺度滤波器对采样的 s 个尺度的测试样本进行缩放, 统一为固定的大小 J 。通过式(3)计算得到 s 个尺度对应的响应值, 取最大响应值对应的尺度作为跟踪框在当前帧图像上需要变化的尺度。

尺度滤波器以目标的位置中心由内向外采样, 提取测试样本的尺度变化是非线性的指数关系, 与上一帧跟踪框的尺寸越接近, 采样就越精细; 尺寸越偏离, 采样就越粗略。DSST 算法可以尺度自适应的应对跟踪过程中目标大小变化的问题。

2 多特征融合

样本的特征表达是影响跟踪算法性能的重要因素^[5]。DSST 算法采用目标的 HOG 特征。HOG 特征很好地描述目标的纹理特征和轮廓形状, 对颜色变化不敏感, 在色彩复杂、忽明忽暗的背景下也能很好地跟踪目标。HOG 特征在目标形变和快速运动的情况下效果会变差。颜色特征能够帮助跟踪算法解决运动模糊问题, 但对于色彩变化敏感。灰度增强对像素灰度值进行处理, 增强图像对比度。LBP 特征描述图像的局部纹理特征, 将一个像素点灰度值与全局灰度值的关系转变为中心像素点与其邻域像素灰度值的关系, 在图像中明暗变化明显的分界处能够体现边缘信息, 而且根据灰度不变性来应对光照变化。笔者将增强的灰度特征、HOG 特征、LBP 特征、Lab 颜色空间融合到一起, 起到互补的优势, 算法中不同的相关滤波器选择不同特征的组合。



(e) 等价模式 LBP 可视化图 (f) Lab 颜色空间可视化图

图 1 图像特征可视化

2.1 增强的灰度特征

图像的灰度范围是 $[0, 255]$ ，设其灰度级范围是 $[0, L-1]$ ，其灰度直方图表示：

$$h(r_k)=n_k, k=1, 2, \dots, L-1. \quad (6)$$

式中： r_k 是第 k 级的灰度值； n_k 是灰度值为 r_k 的像素点个数； $h(r_k)$ 是以灰度级为横坐标、各灰度级对应的像素个数为纵坐标构成的直方图。

对直方图进行归一化处理：

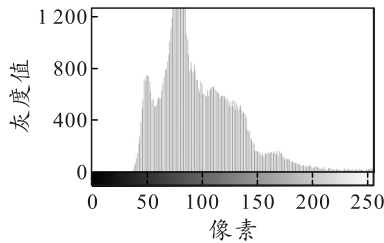
$$p(r_k)=n_k/MN, k=1, 2, \dots, L-1. \quad (7)$$

式中： M 、 N 是图像横纵方向像素总数； $p(r_k)$ 表示各灰度级对应的像素频率。

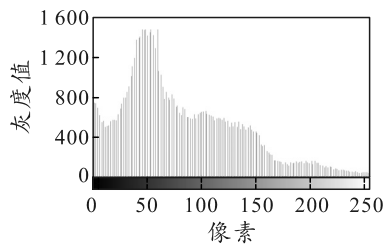
图 2(a)是图 1(b)的灰色直方图，可以发现图像的灰度值一般集中在 $50 \sim 200$ 之间。通过将灰度值均匀地分布在 $0 \sim 255$ 之间，并且将小于 50 的灰度值变为 0 ，大于 200 的灰度值变为 255 ，这样就会使图像变得更加清晰。设 x 为原图像的灰度值， y 为增强图像的灰度值，进行变换如下：

$$(x-50)/(200-x)=(y-0)/(255-y). \quad (8)$$

图 2(b)为灰度变换增强后图像的灰度直方图，图 1(c)为灰度变换增强后的图像。设灰度级为 8 ，将灰度变换增强后图像目标区域的每一个像素点灰度值减去目标区域的平均灰度值，得到目标区域增强的灰度特征。



(a) 灰度直方图



(b) 灰度变换增强的灰度直方图

图 2 灰度变换增强前后的灰度直方图

2.2 HOG 特征

HOG 特征计算图像像素点的方向梯度信息，HOG 特征是基于灰度图处理。

首先将图 1(a)变换为灰度图(b)：

$$I=0.299R+0.587G+0.114B. \quad (9)$$

然后通过 GAMMA 校正去除图像的噪声和光照干扰：

$$I(x,y)=I(x,y)^{\text{gamma}}. \quad (10)$$

式中 gamma 取 0.3 。

计算图像中像素点的梯度：

$$g_x=I(x+1,y)-I(x-1,y); g_y=I(x,y+1)-I(x,y-1). \quad (11)$$

式中 g_x, g_y 分别是水平和垂直方向的梯度。

求该像素点的梯度幅值和梯度方向：

$$\theta(x,y)=\tan^{-1}(g_y/g_x); A(x,y)=\sqrt{(g_x^2+g_y^2)}. \quad (12)$$

最后将图像的局部像素点进行划分，一个 cell 包含 4×4 的像素点。以梯度方向为横轴，以梯度幅值为纵轴，构成 9 个 bin 的直方图来描述每个 cell 的 HOG 特征。每个 bin 的步长是 20° 的方向划分，计算的 HOG 特征包含了 27 维的 bin 通道特征和 4 维的反映当前 cell 周围区域能量的梯度能量特征。其中 27 维的 bin 通道特征包含 9 个对方向不敏感的 bin 通道 (20° 的步长对 $0 \sim 180^\circ$ 的方向划分) 和 18 个对方向敏感的 bin 通道 (20° 的步长对 $0 \sim 360^\circ$ 的方向划分)；因此，图像中一个像素点的 HOG 特征包含 31 维通道的特征向量。

2.3 LBP 特征

LBP 特征计算灰度图的局部二值信息如图 3 所示，采用一个 3×3 的窗口，对窗口中心的像素点进行处理：

$$\text{LBP}(x_c, y_c) = \sum_{p=0}^{p-1} 2^p s(i_p - i_c). \quad (13)$$

式中： (x_c, y_c) 是中心像素点； i_c 是其灰度值； i_p 是周围邻域的灰度值； s 是二值函数，如果 $i_p < i_c, s=0$ ；如果 $i_p \geq i_c, s=1$ 。计算得到一个像素点对应的 8 位 2 进制数，从左上角沿顺时针方向取值，转为 10 进制后得到该像素点对应的 LBP 特征。

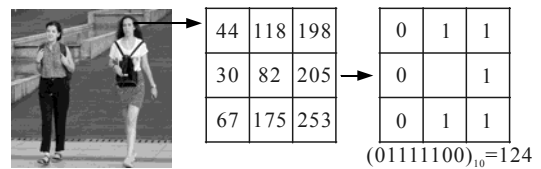


图 3 LBP 特征的提取

将 3×3 大小的正方形邻域变为任意大小的圆形邻域来改进 LBP 算子，以适应不同尺寸的纹理特征，还实现了特征旋转不变性。

对于 3×3 邻域的 LBP 算子，有 2^8 种二进制模式，导致直方图表示过于稀疏，而且统计量过于复杂。利用等价模式 (uniform pattern) 对 LBP 算子的

模式进行降维。模式从 255 种降到了 $p \cdot (p-1) + 2$ 种， p 为邻域中采样点数，图 2 中 $p=8$ 。

2.4 Lab 特征

Lab 颜色空间又称为 CIELab，Lab 由 3 个通道构成，其中 1 个 L 亮度通道和 2 个 a、b 颜色通道，L 通道表示像素点的亮度，a 通道表示从红色到绿色的范围，b 通道表示从黄色到蓝色的范围。Lab 相较于 RGB，色域更宽，符合人类的视觉感知，也更容易调节，Lab 特征想要调节亮度就只需要改变 L 通道，想要调节颜色就调节 a 和 b 通道，而 RGB 特征需要 R、G、B 3 个通道共同调节。

RGB 不能直接转换为 Lab。

首先，将 RGB 转换为 XYZ 颜色空间：

$$\begin{aligned} X &= 0.4124R + 0.3576G + 0.1805B, \\ Y &= 0.2126R + 0.7152G + 0.0722B, \\ Z &= 0.0193R + 0.1192G + 0.9505B. \end{aligned} \quad (14)$$

然后，将 XYZ 颜色空间转换为 Lab：

$$\begin{aligned} L &= 116f(Y/Y_n) - 16, \\ a &= 500[f(X/X_n) - f(Y/Y_n)], \\ b &= 200[f(Y/Y_n) - f(Z/Z_n)]. \end{aligned} \quad (15)$$

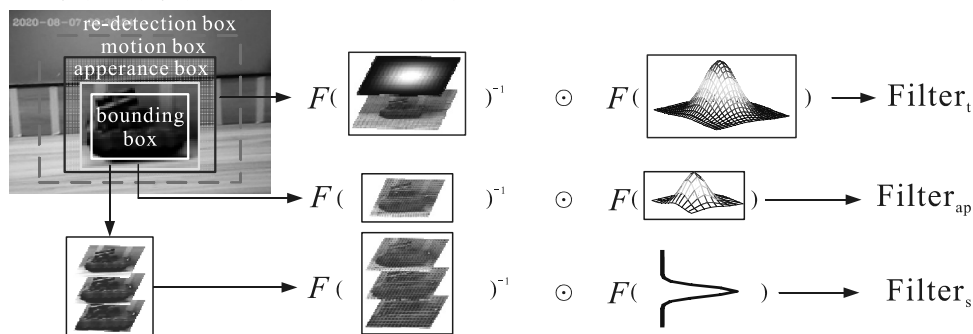


图 4 相关滤波器的训练过程

上图中目标区域由内而外的框依次是目标跟踪框 (bounding box)、外观窗口 (appearance box)、位置窗口 (motion box)、重检测窗口 (re-detection box)。

跟踪框代表当前帧图像中目标的位置和大小。

外观窗口是以跟踪框为中心，跟踪框上下左右分别扩大 4 个像素的区域。利用外观窗口提取 Fhog 特征、LBP 特征和增强的灰度特征，并且将特征融合来训练外观滤波器 $Filter_{ap}$ ；同时外观窗口通过建立尺度池中的缩放因子得到 33 个不同大小的图像块，只采样这些图像块的 Fhog 特征，统一尺寸为 J 后，训练尺度滤波器 $Filter_s$ 。

位置窗口是以跟踪框为中心，跟踪框大小 2.8 倍的区域。在位置窗口中提取图像块的 Fhog 特征、LBP 特征和增强的灰度特征，将特征融合来训练位

其中： $f(t) = \begin{cases} t^{1/3} & \text{if } t > 0.0089 \\ 7.7091t + 0.1379 & \text{otherwise} \end{cases}$ ； X_n 、 Y_n 、

Z_n 取值为 95.047、100.000、108.883。彩色图像一个像素点的 Lab 特征对应一个 3 维的特征向量，图 1(f) 是 Lab 特征可视化的图像。

2.5 特征的矢量融合

笔者将上文介绍的各个特征通过矢量相加的方法进行融合^[6]。

$$K^{xx'} = \exp\left(-\frac{1}{\sigma^2}(\|x\|^2 + \|x'\|^2) - 2F^{-1}(\hat{x}_c \otimes \hat{x}_c')\right). \quad (16)$$

式中： x 是单一特征； x_c 是多特征矢量相加后的融合特征。

3 融合 SVM 的 DSST 算法

3.1 相关滤波器的训练

笔者在 DSST 算法位置滤波器和尺度滤波器基础上，增加一个外观滤波器。如图 4 所示，在图像上采用不同大小的窗口采样来训练不同的滤波器。首帧图像人工标注目标跟踪框 (x_0, y_0, w_0, h_0) ，标注的跟踪框表示目标样本—图像中的遥控坦克。

置滤波器 $Filter_t$ 。位置窗口采样得到的样本包含目标 (正样本) 和背景 (负样本)，用正、负样本来训练滤波器，增强了位置滤波器对目标的判别能力。

重检测窗口是位置窗口再扩大 1.2 倍的区域，当外观滤波器判定目标丢失后^[7]，SVM 重检测器被激活，在重检测窗口内搜索目标。

位置滤波器在位置窗口内采样时加入汉宁窗，汉宁窗减弱了训练样本循环矩阵造成的边界效应问题。位置和外观滤波器在训练时结合了不同大小的 2 维高斯标签，尺度滤波器采用了 1 维高斯标签。高斯标签赋予靠近图像块中心样本更高的权重，远离中心样本更低的权重，使得样本并非传统意义上的非正即负，高斯加权的方式在实际采样中的效果更好。

3.2 相关滤波器的更新

位置和尺度滤波器采用传统的模型更新方式，对每帧图像采样来更新模型。根据式(4)，结合历史模型参数，用线性加权对模型更新，避免更新的剧烈程度，在一定程度上减少噪声对模型的污染。

外观滤波器采用稀疏更新的方式。根据式(3)，只有当外观滤波器计算的外观响应值的最大值高于设定的外观阈值时，即

$$\max_res_{ap} > \theta_{ap}。 \quad (17)$$

当前帧图像的外观窗口提取样本的融合特征来更新外观模型，同样使用线性加权的方法；如果 $\max_res_{ap} < \theta_{ap}$ ，当前帧就不会更新外观模型。稀疏更新的方式使外观滤波器保留了可靠的样本学习模型，避免了在目标被遮挡情况下仍然更新对模型造成的污染。

3.3 目标重检测

3.3.1 SVM 分类器概述

SVM 是一种有监督的机器学习，目的是建立一个最优分类面(超平面)，超平面表示如下：

$$w \cdot x + b = 0。 \quad (18)$$

构建的超平面使 2 类点集的边缘点(支持向量)到此超平面的距离达到最大，实现分类的目的。实质上，训练 SVM 的过程就是不断引入支持向量求解最优分类面的参数 w 和 b 的过程。SVM 引入了径向基核函数和松弛变量来解决 2 类点集的线性不可分问题。由于在图像的目标样本附近密集采样，采样的样本数目巨大，但 SVM 仅利用那些作为“支持向量”的样本来训练分类面，极大简化了计算。

3.3.2 SVM 分类器的训练

对于某一帧图像，位置和尺度滤波器分别估计出目标在图像中的位置和尺度，得到当前帧的跟踪框。根据式(17)，只有当外观滤波器计算的最大响应值大于 θ_{ap} 时，外观滤波器才进行更新，否则不更新外观滤波器。同时，在重检测窗口中对当前帧图像提取样本特征来训练 SVM 的参数。

样本特征用 1 通道的增强灰度特征和 3 通道的 Lab 特征组成的 4 通道特征表示，用直方图表示测试样本的特征向量。

计算样本集中每个测试样本与目标样本的交并比，得到每个测试样本与目标样本的重叠度：

$$IOU = (\text{area}(x_i) \cap \text{area}(x_j)) / (\text{area}(x_i) \cup \text{area}(x_j))。 \quad (19)$$

式中： x_i 表示训练集中的每个测试样本； x_j 表示目

标样本。

在样本集中，与目标样本的交并比 $IOU < 0.5$ 的测试样本作为负样本，设置其标签值是 -1， $IOU > 0.5$ 的测试样本作为正样本，设置其标签值是 1，其余样本删除，只有正、负样本才作为训练 SVM 的支持向量。假设在图像重检测窗口采集 N 个测试样本，样本集表示为 $\{(v_i, y_i) | i = 1, 2, \dots, N\}$ ，其中， v_i 表示第 i 个样本 x_i 的特征向量， y_i 表示第 i 个样本标签，取值 ± 1 。用样本集训练 SVM， w 的目标方程是：

$$\min_w (\lambda/2) \|w\|^2 + (1/N) \sum_i l(w; (y_i, c_i))。 \quad (20)$$

式中： $l(w; (v_i, y_i)) = \max\{0, 1 - y_i \cdot \langle w, v_i \rangle\}$ ；实验设 λ 大小是 0.000 1。

3.3.3 SVM 分类器对目标的重检测

对新的一帧图像，位置滤波器计算响应值估计目标的位置。外观滤波器以此为中心，在外观窗口中提取测试样本，根据式(3)来计算其外观响应值 res_{ap} 。当外观响应值的最大值 \max_res_{ap} 低于位置阈值 θ_t 时，即 $\max_res_{ap} < \theta_t$ ，外观滤波器就判定目标跟踪失败，此时激活 SVM^[8]。

SVM 在当前帧图像的绿色窗口范围内搜索目标，采样获得样本集合 $\{v_i | i = 1, 2, \dots, N\}$ ，其中 v_i 是每个样本的特征向量。SVM 对每个样本打分：

$$score_i = w \cdot v_i + b, i = 1, 2, \dots, N。 \quad (21)$$

式中： w 和 b 是 SVM 在之前帧图像中训练好的最优分类面参数； v_i 是当前帧绿色窗口采样的每个样本的特征向量； $score_i$ 是 SVM 对每个样本的打分。

取最大得分的样本作为 SVM 假设的正样本，对外观滤波器的正样本进行二次判定，计算其外观响应值 SVM_res_{ap} 。如果 SVM 分类后的正样本的外观响应值最大值 $\max_SVM_res_{ap}$ 大于跟踪失败时位置滤波器推导的目标样本的外观响应值最大值 \max_res_{ap} 的 1.5 倍，即：

$$\max_SVM_res_{ap} \geq 1.5 \max_res_{ap}。 \quad (22)$$

外观滤波器就认为 SVM 分类后的这个正样本是目标，这样在目标丢失后就能重检测到目标。如果不满足式(22)的关系，说明 SVM 对当前帧图像的目标重检测失败，仍旧把跟踪失败时位置滤波器推导的目标样本当作目标。

3.3.4 目标重检测的跟踪流程

DSST 算法结合目标重检测的跟踪流程如图 5 所示。

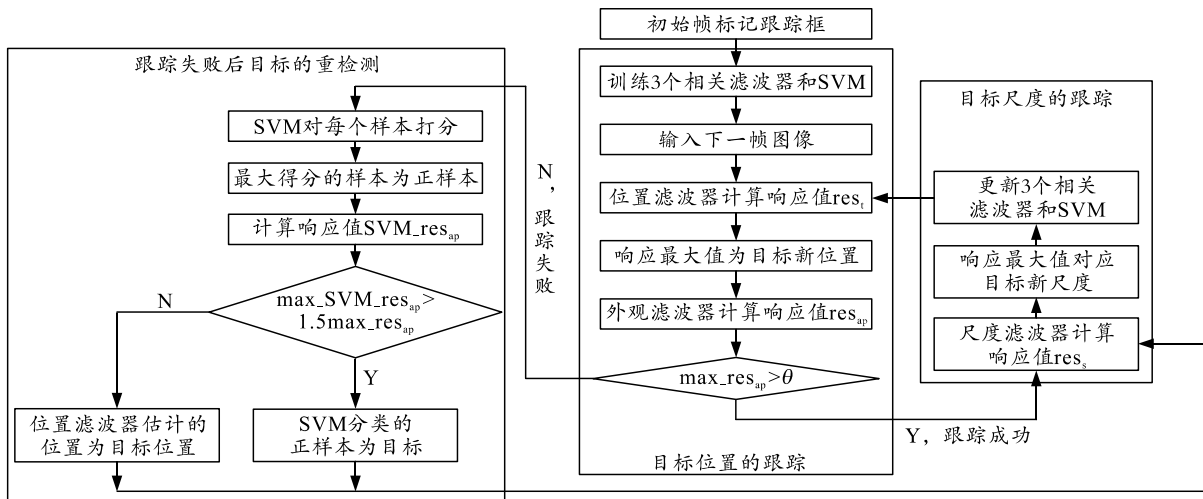


图 5 DSST 算法结合目标重检测的跟踪流程

4 实验结果与仿真分析

4.1 算法对目标的跟踪与分析

实验中电脑 CPU 配置是 Intel Core i5-8033H 四核处理器。改进算法与 DSST 算法的参数基本保持一致。在尺度估计模块中设置 $a=1.02$ ，尺度个数 $s=33$ 。重检测模块的位置阈值 $\theta_r=0.15$ ，外观阈值 $\theta_{ap}=0.38$ 。跟踪过程中 KCF 算法用点线框表示，DSST 算法用虚线框表示，融合重检测的 DSST 算法用实线框表示。

图 6 是用红外热像仪拍摄的一段无人机夜间飞行视频截图，飞行中目标发生旋转，用 3 种算法跟踪。开始 3 种算法对目标跟踪良好。第 27 帧中由于无人机旋转带来的目标外观变化，3 个跟踪框对无人机的跟踪产生偏差，KCF 算法和 DSST 算法仍然采样更新模型，模型中误差不断积累，改进算法的外观滤波器判断出目标丢失，激活 SVM。第 28 帧改进算法的 SVM 全局搜索目标，SVM 分类的正样本被外观滤波器二次判定后认为是目标，这样改进算法在第 28 帧中重检测到目标，KCF 算法和 DSST 算法跟丢目标导致跟踪失败。

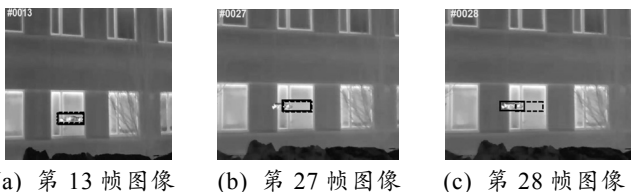


图 6 3 种算法对旋转飞行的红外目标跟踪

图 7 是拍摄的一段玩具大小发生变化的可见光视频截图，用 3 种算法跟踪目标的“头部”。在目标大小不发生改变的情况下，如第 391 帧，3 个跟踪框都对玩具的“头部”跟踪良好，但是在目标变大

或变小的过程中，如第 899 帧和第 1323 帧，DSST 算法和改进算法都自适应地改变跟踪框尺寸，但是 KCF 算法的跟踪框的尺寸始终不会调整。



图 7 3 种算法对大小变化的可见光目标跟踪

图 8 是拍摄的一段人体脸部发生的移动、旋转、被其他人脸遮挡的可见光视频截图，用 3 种算法跟踪目标。由图 8(a)–(c)所示，在目标人脸移动旋转过程中，3 种算法对目标跟踪良好，图 8(d)表示目标被严重遮挡，此时 KCF 算法和 DSST 算法的模型被遮挡物体污染，此后将遮挡物体看成是目标。改进算法判断出目标丢失，调用 SVM 对目标进行重新检测，SVM 重检测的正样本由外观滤波器二次判定后认为是目标，改进算法对目标的重检测成功，图 8(e)表示改进算法能够准确跟踪目标人脸。

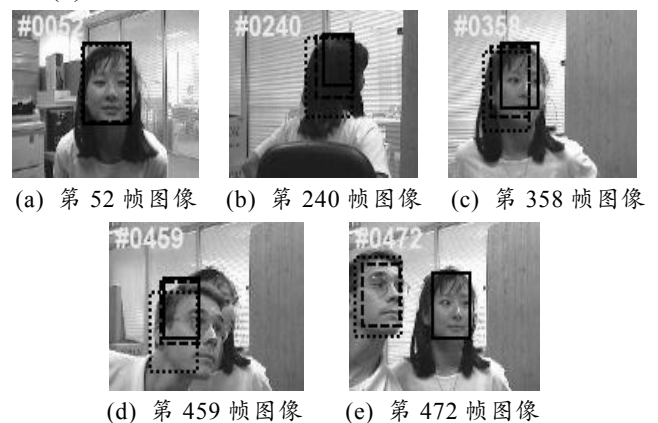


图 8 3 种算法对受到严重遮挡的目标跟踪

4.2 3 种算法在 OTB-2013 上的测试对比

OTB-2013(object tracking benchmark) 根据算法在平台上测试的成功率和精确度来定量地评估算法的跟踪性能。它统一了各类算法的输入输出格式, 并且建立一个由人工标注目标真值(groundtruth)的视频序列集。序列集包含 50 个视频, 代表了不同的跟踪难点问题, 有遮挡、尺度变化等 11 类跟踪难点问题。

将 KCF 算法、DSST 算法、改进算法在 OTB-2013 上测试, 得到算法的精确度和成功率图, 定量地评估算法的跟踪性能。精确度是基于算法运行的跟踪框与每帧图像人工标注好的跟踪框真值的中心位置误差; 成功率是基于算法运行的跟踪框与人工标注的跟踪框真值的重叠率误差。上述测试是算法的一次性评估(one-pass evaluation, OPE)。

如图 9 所示, 改进算法在一次性评估结果的精

确度达到了 0.832, 成功率也达到了 0.805。相比于 KCF 算法和 DSST 算法, 改进算法很大程度地提高了目标跟踪的性能。由于篇幅有限, 笔者只列举了遮挡和目标尺度变化问题上的测试结果。在目标遮挡问题上, DSST 和 KCF 算法都没有采用有效的策略来应对目标受到遮挡时的跟踪, 但改进算法对增强的灰度特征、HOG 特征、LBP 特征、Lab 特征的融合, 丰富了目标的特征表达, 使改进算法从背景信息中对目标的判别能力更好, 而且增加的外观相关滤波器可以有效地判定目标受到遮挡时跟踪是否失败, 融合的 SVM 对跟丢的目标再全局搜索, 直到重检测到目标, 所以改进算法对遮挡目标跟踪的效果要远远好于 DSST 和 KCF 算法。在目标尺度变化问题上, KCF 算法采用大小不变的跟踪框对目标跟踪, 而改进算法和 DSST 算法都建立了尺度池, 可以自适应地变尺度跟踪, 二者跟踪精度相近, 都远远超过了 KCF 算法。

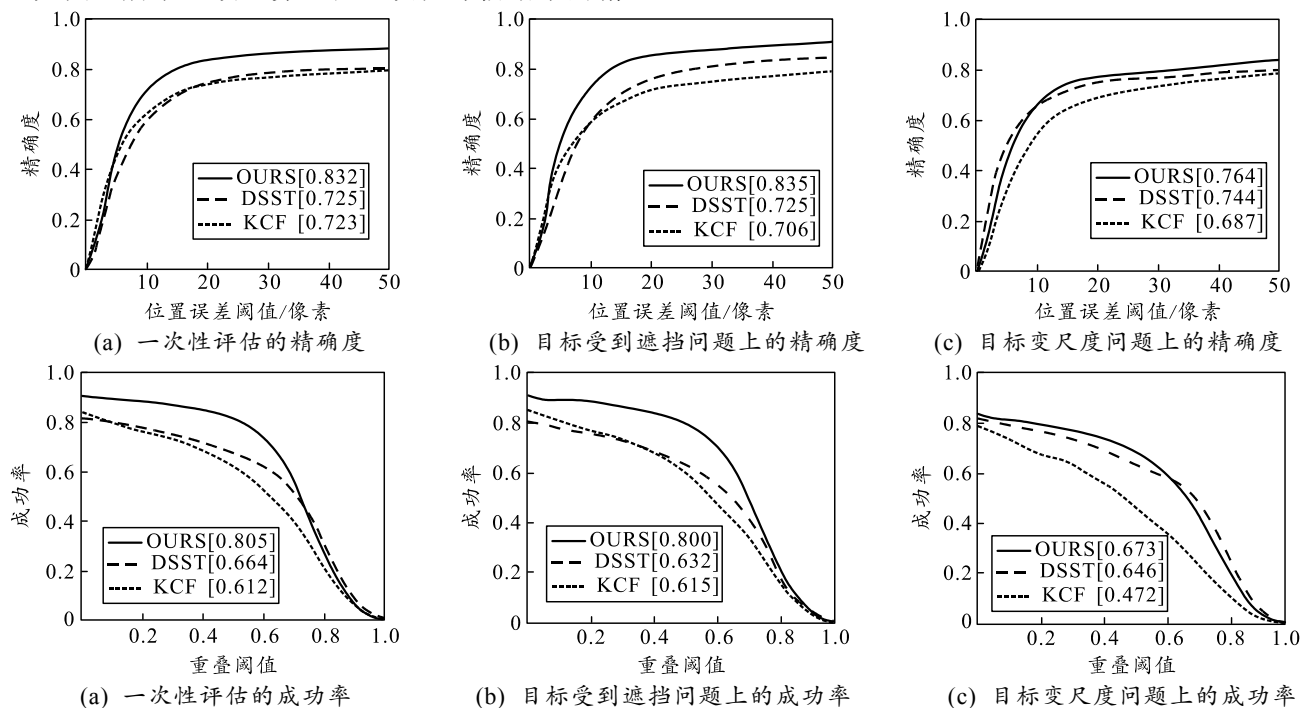


图 9 3 种算法在 OTB-2013 测试的精确度和成功率

5 结论

目标跟踪技术已经取得了显著进步, 在军事目标侦察和导弹制导上的应用越来越广泛。相关滤波算法仍需要进一步发展和完善, 笔者在 DSST 算法的基础上, 将多特征进行了融合, 并在位置和尺度滤波器的基础上, 建立了目标外观滤波器, 而且将 SVM 目标重检测模块融合进算法中, 使算法在应对目标大小变化、目标形变、目标旋转、物体遮挡等

问题上跟踪鲁棒性有了明显提升。

参考文献:

- [1] BOLME D S, BEVERIDGE J R, DRAPER B A, et al. Visual object tracking using adaptive correlation filters[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. San Francisco, California, United States: IEEE, 2010: 2544-2550. [DOI: 10.1109/CVPR.2010.5539960].