

基于全局背景与特征降维的视觉跟踪算法

孙彦景* 王赛楠 石韞开 云霄 施文娟

(中国矿业大学信息与控制工程学院 徐州 221116)

摘要: 相关滤波算法容易受到形变、运动模糊、相似背景等因素的干扰,导致跟踪任务失败。为了克服以上问题,该文提出一种基于全局背景与特征降维的视觉跟踪算法。该算法首先提取紧邻目标的图像区域作为负样本供分类器学习,以抑制相似背景的干扰;然后提出一种基于主成分分析的更新策略,构建降维矩阵压缩HOG特征的维度,在更新分类器的同时减少其冗余度;最后加入颜色特征表征运动目标,并根据特征对系统状态的响应强度进行自适应融合。在标准数据集上将该文提出的算法与Staple, KCF等其他算法进行了仿真对比,结果表明该文算法具有更强的鲁棒性,在形变因素的影响下,所提出的算法与Staple和KCF算法相比距离精度分别提升8.3%和13.1%。

关键词: 视觉跟踪; 全局背景信息; 特征降维; 自适应融合

中图分类号: TP391.4

文献标识码: A

文章编号: 1009-5896(2018)09-2135-08

DOI: 10.11999/JEIT171143

Visual Tracking Algorithm Based on Global Context and Feature Dimensionality Reduction

SUN Yanjing WANG Sainan SHI Yunkai YUN Xiao SHI Wenjuan

(School of Information and Control Engineering, China University of Mining Technology, Xuzhou 221116, China)

Abstract: Tracking effects of algorithms using correlation filter are easily interfered by deformation, motion blur and background clustering, which can result in tracking failure. To solve these problems, a visual tracking algorithm based on global context and feature dimensionality reduction is proposed. Firstly, the image patches uniformly around the target are extracted as negative sample, and thus the similar background patches around the target are suppressed. Then, an update strategy based on principal component analysis is proposed, constructing the matrix to reduce the dimensionality of HOG feature, which can reduce the redundancy of feature when it updates. Finally, the color features are added to represent the motion target and the response of the system states are adaptively fused according to the features. Experiments are performed on recent online tracking benchmark. The results show that the proposed method performs favorably both in terms of accuracy and robustness compared to the state-of-the-art trackers such as Staple or KCF. When deformation occur, the proposed method is shown to outperform the Staple tracker and KCF algorithm by 8.3% and 13.1% respectively in median distance precision.

Key words: Visual tracking; Global context information; Feature dimensionality reduction; Adaptive fusion

收稿日期: 2017-12-04; 改回日期: 2018-05-02; 网络出版: 2018-07-12

*通信作者: 孙彦景 yjsun@cumt.edu.cn

基金项目: 江苏省自然科学基金青年基金(BK20150204), 国家重点研发计划(2016YFC0801403), 国家自然科学基金(51504214, 51504255, 51734009, 61771417), 江苏省重点研发计划(BE2015040)

Foundation Items: The Natural Science Foundation of Jiangsu Province (BK20150204), The State Key Research Development Program (2016YFC0801403), The National Natural Science Foundation of China (51504214, 51504255, 51734009, 61771417), The Research Development Programme of Jiangsu Province (BE2015040)

1 引言

视觉目标跟踪是在视频序列中根据第1帧的目标状态信息估计其后续状态的过程^[1],在人机交互、视频监控和机器人等领域具有广泛的应用。近年涌现大量的目标跟踪算法,但由于形变、运动模糊和快速运动等因素影响,目标跟踪仍是计算机视觉领域的难点。

根据目标外观模型,跟踪算法分为生成式模型^[2-4]和判别式模型^[5-7]。判别式模型通过提取正负样本特征,将跟踪问题转化为二分类问题,寻找最佳判别界限区分目标和背景,基于判别式模型的目标跟踪算法是近年的研究重点^[8],如基于核结构化输出(Structured output tracking with kernel, Struck)^[5]跟踪算法、跟踪-学习-检测(Tracking-Learning-Detection, TLD)^[6]跟踪算法、多样例学习(Multiple Instance Learning, MIL)^[7]跟踪算法等。但是这些算法需要大量的样本,且物体在运动过程中存在复杂的影响因素,包括光照变化、部分遮挡、尺度变化、运动模糊等,这些都制约着目标跟踪算法的性能。

相关滤波目标跟踪算法是基于判别式模型目标跟踪算法的重要理论分支,其过程可描述为在训练阶段学习滤波模板,在检测阶段查找相关响应图的峰值来定位目标。Bolme等人^[9]首次将相关滤波应用于目标跟踪中,提出了MOSSE(Minimum Output Sum of Squared Error)跟踪算法。Henrique等人^[10]提出显著性相关滤波,该算法利用循环移位样本近似密集采样,大大地增加了训练样本的数目,提升了分类器的辨别能力,随后又将核技术融入算法,提出了核化相关滤波(Kernelized Correlation Filter, KCF)^[11]算法。Danelljan等人^[12]提出基于相关滤波的方法训练尺度滤波器,以适应跟踪过程中尺度变化。上述基于相关滤波的目标跟踪算法与传统算法相比性能较好,但算法可学习的背景信息有限,导致分类能力低。Mueller等人^[13]对MOSSE跟踪算法的分类器进行了改进,提出一种上下文感知目标跟踪算法,达到了鲁棒估计的效果,使用HOG特征描述目标外观,由于冗余导致计算复杂。Staple算法^[14]将颜色特征和模板类特征(HOG特征)结合,以更好地表征目标外观,使用固定系数进行特征融合,没有考虑不同特征的重要性。

综合以上基于相关滤波的目标跟踪算法可以看出:(1)相关方法仅关注目标信息,而忽略了背景信息学习的重要性。因此当目标发生形变或者遮挡时,容易引入背景造成跟踪丢失;(2)相关方法使用HOG特征描述目标外观模型时,高维特征导致

计算复杂,降低跟踪效率;(3)使用固定的系数进行特征融合,无法应对复杂的视频环境。

基于上述讨论本文提出了一种基于全局背景与特征降维的目标跟踪算法。所提算法提取紧邻目标的图像区域作为负样本供分类器学习,充分利用了目标周围的背景信息,从而避免了形变情况下背景中相似物的干扰;在更新模板滤波器模型参数时,引入了降维的策略,构建PCA(Principle Component Analysis)降维矩阵压缩HOG特征维度以减少运算量提升速度;最后在响应图层面根据特征重要性调整系数,达到自适应融合的目的。实验结果表明,本文提出的跟踪算法能准确可靠地跟踪目标,并很好地处理形变、背景干扰以及运动模糊的影响。

2 相关滤波目标跟踪算法

相关滤波算法选择HOG特征作为目标模板的描述策略,HOG特征充分利用图像的局部梯度幅值和方向的特征值,可以很好地表征目标的轮廓信息和局部细节特征^[15],对光照因素的变化具有较强的鲁棒性。相关滤波目标跟踪算法主要包括分类器训练、参数更新与目标位置预测,其过程可描述为根据第1帧目标位置提取图像块,训练、学习得到相关滤波器;对于随后的每一帧,从上一帧目标位置的区域提取图像块进行检测,并提取估计位置处的目标外观,训练和更新相关滤波器。

2.1 分类器训练

相关滤波目标跟踪算法采用正则化最小二乘法分类器来预测目标位置,对所有的训练样本 \mathbf{f} 和期望输出 \mathbf{g} 通过最小化损失函数 L ,如式(1)所示,求解分类器权重 \mathbf{h} ,损失函数定义为

$$L(\mathbf{h}) = \|\mathbf{g} - \mathbf{h} * \mathbf{f}\|^2 + \lambda \|\mathbf{h}\|^2 \quad (1)$$

其中, \mathbf{g} 是以目标为中心的2维高斯函数,代表训练样本 \mathbf{f} 的理想回归值, \mathbf{h} 表示分类器权重, \mathbf{f} 表示基于HOG特征描述的训练样本, λ 是正则化参数,*表示循环卷积。因为时域计算的复杂性,所以利用帕斯瓦尔定理将式(1)转换成频域形式,再根据稠密采样算法^[11]求得分类器权重:

$$\mathbf{H} = \frac{\mathbf{F} \odot \mathbf{G}^*}{\mathbf{F} \odot \mathbf{F}^* + \lambda} \quad (2)$$

其中, \odot 表示点乘操作, \mathbf{F} 和 \mathbf{G} 分别表示 \mathbf{f} 和 \mathbf{g} 的傅里叶正变换, \mathbf{F}^* 是 \mathbf{F} 的复共轭。

2.2 参数更新

在目标跟踪过程中,旋转、形变和光照等因素会造成目标模型变化,需要及时更新分类器以适应目标模型的变化。分类器模型参数更新策略如式(3):

$$\left. \begin{aligned} A_i &= \eta F_i \odot G_i^* + (1 - \eta) A_{i-1} \\ B_i &= \eta F_i \odot F_i^* + (1 - \eta) B_{i-1} \end{aligned} \right\} \quad (3)$$

其中, A_i 和 B_i 分别表示分类器 H_i 的分子和分母的更新方式, η 表示学习速率, i 是当前的帧序号。 $F_i \odot G_i^*$ 和 $F_i \odot F_i^*$ 是训练样本通过式(2)得到的分类器参数。

2.3 目标位置预测

参数更新后的分类器用于预测运动目标的位置, 即采集搜索区域内所有测试样本, 通过检测得到响应输出, 查找最大响应的位置为目标的预测位置。分类器的响应输出如式(4):

$$Y = Z \odot H^* \quad (4)$$

其中, Z 是测试样本, Y 是最终响应输出。

3 基于全局背景信息的实时目标跟踪

针对相关滤波算法存在的相关问题, 首先基于被跟踪目标相邻图像区域的相似性改进了损失函数, 使其在目标外观发生突变或者相似背景干扰时, 仍能准确地跟踪目标; 其次, 提出一种基于主成分分析的更新策略, 构建降维矩阵压缩HOG特征的维度, 使得在更新分类器的同时减少其冗余度; 最后, 设计了一种评价不同特征重要性的标准, 以自适应分配权重达到鲁棒跟踪的目的。

3.1 全局背景提取

相关滤波算法在构建损失函数 $L(h)$ 时只考虑了正则化信息, 因为被跟踪目标的相邻图像区域与其具有一定的相似性, 所以跟踪过程中往往会产生相似物干扰的现象^[13,16], 从而导致跟踪目标框产生漂移。为解决此类问题, 提取紧邻目标的图像区域作为负样本供分类器学习, 以提升分类器的分类能力和鲁棒性。将背景区域的响应值作为正则项加入到相关滤波的损失函数中, 更新损失函数方程如式(5):

$$\begin{aligned} L(h) &= \left\| g - \sum_{l=1}^d h^l * f^l \right\|^2 + \lambda_1 \sum_{l=1}^d \|h^l\|^2 \\ &+ \lambda_2 \sum_{i=1}^k \sum_{l=1}^d \|f_i^l * h^l\|^2 \end{aligned} \quad (5)$$

式中, $\lambda_2 \sum_{i=1}^k \sum_{l=1}^d \|f_i^l * h^l\|^2$ 为新增项, 包含了全局背景信息, 其中 f_i^l 代表第 i 个背景区域提取的第 l 个通道的特征, h^l 代表分类器的第 l 个通道, k 取值为4, 代表紧邻目标的4个背景区域, d 取值为32, 代表32个特征通道, λ_2 是背景区域的约束项, 此处取值0.5。如式(6)对训练样本进行HOG特征提取:

$$f = \phi[u] \quad (6)$$

式中, u 代表训练样本, ϕ 是提取HOG特征的操作, f 为基于HOG特征描述的训练样本。

根据文献^[11]稠密采样算法, 将式(5)变换到傅里叶域求解, 得到分类器训练公式:

$$H^l = \frac{G^* \odot F^l}{\sum_{k=1}^d (F^k)^* \odot F^k + \lambda_1 + \lambda_2 \sum_{i=1}^k \sum_{l=1}^d (F_i^l)^* \odot F_i^l} \quad (7)$$

其中, H 为分类器权重, F 和 G 分别是 f 和 g 的傅里叶正变换, G^* 是 G 的复共轭。利用 A_t^l 和 B_t^l 分别表示 H^l 的分子和分母, 则式(7)变为:

$$H^l = A_t^l / (B_t^l + \lambda_1) \quad (8)$$

其中, λ_1 取值为 e^{-3} , l 表示分类器的第 l 个通道, 因利用HOG特征对样本进行描述, 所以最大取32。

3.2 降维特征的参数更新

由上述可知, 在损失函数中加入背景信息后, 通过优化损失函数求解分类器权重 H , 其中式(7)引入 A_t^l 和 B_t^l 将分类器模型公式简写为式(8), 对模型进行参数更新时因HOG特征之间存在冗余, 所以对HOG特征进行PCA降维, 把降维后的数据用于分类器的参数更新和目标检测, 以提升跟踪的实时性。

3.2.1 HOG特征降维 从样本中提取的HOG特征维度高且各维度间具有较强的相关性^[1,17], 采用主成分分析方法对样本特征压缩降维, 降维过程如式(9):

$$\tilde{U}_t = F \{P_t u_t\} \quad (9)$$

$$C_t = \sum_n u_t(n) u_t^T(n) \quad (10)$$

其中, F 表示矩阵的傅里叶正变换, t 代表当前帧序号。根据协方差矩阵式(10)计算出特征值并按单调递减的顺序排列, 与其对应的特征向量张成一个子空间, 即降维矩阵 P_t 。式(10)中 u_t 表示样本, n 代表样本中的元素, t 是帧序号。

分别对目标样本和训练样本进行降维操作, 得到降维后的训练样本: $\tilde{F}_t = F \{P_t f_t\}$, 以及降维后的目标样本: $\tilde{U}_t = F \{P_t u_t\}$, 用于分类器参数更新, 同时在检测过程中采用上述方法压缩检测样本, 以简化运算提升速度。

3.2.2 参数更新 对目标样本 u_t 和训练样本 f_t 进行PCA降维, 得到降维后的目标样本 \tilde{U}_t 和训练样本 \tilde{F}_t , 用于更新分类器的参数, 其中目标样本 u_t 与训练样本 f_t 之间存在的关系为: $u_t = (1 - \eta) u_{t-1} + \eta f_t$, 表示目标样本的更新策略。

利用降维后的样本更新分类器模型参数, 策略如式(11):

$$\left. \begin{aligned} \tilde{\mathbf{A}}_t^l &= \mathbf{G}^* \tilde{\mathbf{U}}_t^l \\ \tilde{\mathbf{B}}_t &= (1 - \eta) \tilde{\mathbf{B}}_{t-1} \\ &+ \eta \left\{ \sum_{l=1}^{\tilde{d}} (\tilde{\mathbf{F}}_t^l)^* \tilde{\mathbf{F}}_t^l + \lambda_2 \sum_{n=1}^k \sum_{l=1}^{\tilde{d}} [\tilde{\mathbf{F}}_t^l(n)]^* \tilde{\mathbf{F}}_t^l(n) \right\} \end{aligned} \right\} \quad (11)$$

式中, $\tilde{\mathbf{A}}_t^l$ 和 $\tilde{\mathbf{B}}_t$ 分别表示分类器分子和分母的更新策略, $\tilde{\mathbf{A}}_t^l$ 表示当前帧的分子, $\tilde{\mathbf{B}}_t$ 和 $\tilde{\mathbf{B}}_{t-1}$ 分别表示当前帧和上一帧的分母, t 代表当前帧的序号, η 是学习速率, 取值为0.015, $\tilde{\mathbf{F}}_t^l(n)$ 表示降维后的背景样本, n 表示第 n 个背景块, k 取值为4, 表示选取紧邻目标的4个背景块, \tilde{d} 是降维后的HOG特征维度, 取值为17。

更新后的分类器参数 $\tilde{\mathbf{A}}_t^l$ 和 $\tilde{\mathbf{B}}_t$ 用于下一帧的检测:

$$\mathbf{Y}_{t+1} = \frac{\sum_{l=1}^{\tilde{d}} (\tilde{\mathbf{A}}_t^l)^* \tilde{\mathbf{Z}}_t^l}{\tilde{\mathbf{B}}_t + \lambda_1} \quad (12)$$

其中, \mathbf{Y}_{t+1} 表示基于HOG特征的滤波响应图, $\tilde{\mathbf{Z}}$ 是降维后的检测样本。

3.3 自适应特征融合的目标位置估计

相关滤波利用HOG特征描述目标的轮廓, 但是单一特征对目标快速运动以及形变敏感^[18], 因此, 在原HOG特征基础上, 加入颜色特征表征运动目标, 并根据特征重要性进行自适应融合, 以提升目标跟踪的准确性, 本文采用统计直方图表征目标的颜色特征^[14]。

为满足不同视频序列对特征的需求, 本文根据特征重要性进行权重的更新, 特征越重要即该特征下目标跟踪的效果越好, 在下一帧对其赋更高的权

重。通过式(13)衡量特征重要性:

$$\gamma_i = \frac{1/(f_i - 1)^2}{1/(f_i - 1)^2 + 1/(f_j - 1)^2} \quad (13)$$

两种特征代表的内容和意义不同, 因此在响应图层面将颜色特征和HOG特征进行自适应融合, 如式(14):

$$f(x) = \gamma_t f_t(x) + \gamma_h f_h(x) \quad (14)$$

其中, $f_t(x)$ 为HOG特征响应输出值, 是利用分类器 \mathbf{H} 以及式(12)得到的响应值, $f_h(x)$ 为颜色特征响应输出值, 是直方图滤波器通过检测得到的响应值^[14], 其中 γ_t 和 γ_h 初始值设为0.7和0.3, 并根据特征的重要性不断更新。

查找融合后响应图中响应值最大的位置, 确定为目标位置。

3.4 算法流程

综合上述3节给出了一种基于全局背景与特征降维的视觉跟踪算法, 算法流程见表1。其中, 目标的初始位置需要手动标定。

4 实验结果与分析

为验证本文算法的性能, 在标准数据集OTB^[19]上将本文算法和经典算法KCF, TLD, Struck以及最新的跟踪算法Staple进行了比较。试验中通过比较算法的距离精度和成功率来评估算法性能: 距离精度是指视频中预测目标位置与实际目标位置之间的中心误差小于阈值的帧数占总帧数的百分比; 成功率是指视频中预测目标位置与实际目标位置之间的重叠率大于阈值的帧数占总帧数的百分比。

4.1 实验参数设置

实验硬件条件为: PC机, Windows10系统, Intel(R)Core(TM)i5-4590CPU, 算法实验平台为Matlab R2016a。测试数据选用标准跟踪数据集

表1 基于全局背景与特征降维的视觉跟踪算法

输入: 图像 I_1, I_1, \dots, I_T , 目标的初始位置 x_1 。

输出: 每帧图像中目标的位置 x_t 。

(1)For $t=1, 2, \dots, T-1$ do:

(2)根据输入图像 I_t 和目标位置 x_t 进行全局背景提取, 利用HOG特征表征目标以及背景, 获得基于HOG特征的训练样本; 同时采集图像块并提取颜色特征;

(3)构建降维矩阵压缩HOG特征, 并利用式(11)对分类器模型进行参数更新;

(4)利用颜色特征表征的样本训练更新直方图分类器;

(5)根据输入图像 I_{t+1} 采集检测样本, 并分别提取HOG和颜色特征; 然后根据式(10)构建降维矩阵, 以压缩32维HOG特征至17维, 得到降维后的检测样本;

(6)将HOG和颜色特征表征的检测样本送入相应的分类器, 由式(12)计算得到基于HOG特征的响应图 f_t ; 利用直方图滤波器进行检测, 获得基于颜色特征的响应图 f_h ;

(7)根据式(13)计算特征权重 γ_i ; 然后利用式(14)将 f_t 和 f_h 自适应融合, 得到最终响应图 f ; 最后查找响应图 f 的峰值来确定目标位置 x_{t+1} 。

(8)End for

(Object Tracking Benchmark, OTB)中的51个完全标注的视频数据集, 包含了11种视频中常见的挑战场景: 光照变化、尺度变化、遮挡、形变、运动模糊、快速运动、平面内旋转、平面外旋转、超出视野、背景杂波、低分辨率。评价指标选取一次性评估^[19](One-Pass Evaluation, OPE), 用ground-truth中目标的位置初始化第1帧, 然后运行跟踪算法得到平均精度和成功率。测试代码以及测试数据来自Visual Tracker Benchmark^[19]。

实验中使用的背景区域数目为4个, HOG特征维度压缩至 $\tilde{d}=17$, λ_1 取 e^{-3} , λ_2 取0.5, 式(11)中 η 取0.015, 式(14)中 γ_t 和 γ_h 初始值设为0.7和0.3。

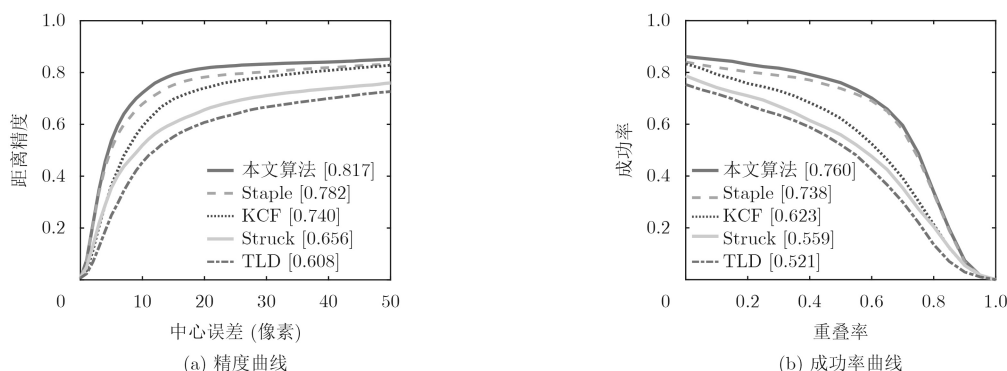


图1 总体的精度曲线和成功率曲线图

法距离精度值达0.817, 相比第2名Staple和第3名KCF分别提升了3.5%和7.7%; 成功率值达0.760, 相比第2名Staple和第3名KCF分别提升了2.2%和13.7%。

(2)基于视频集属性的分析: 利用OTB标准库的11个属性对5种跟踪算法进一步比较分析, 图2依次展现了在形变、运动模糊、快速运动和片外旋转因素响应下的跟踪效果。

通过对图2进行分析, 可以看出本文算法在形变、运动模糊、快速运动和片外旋转等因素影响下均具优势。在形变因素的影响下, 本文提出的算法距离精确度DP为0.871, 相比于Staple和KCF算法分别提高了8.3%和13.1%; 成功重叠率OS为0.818, 相比于Staple和KCF算法分别提高了4.4%和14.7%。表明本文算法的抑制背景策略与自适应特征融合取得了良好的效果。

4.3 定性比较

本文从OTB标准库中的51个视频序列中选取了6个包含多种属性的视频集, 如表2所示, 使用带有多种属性的视频集对本文算法进行定性分析, 并对比本文算法、Staple算法和KCF算法的跟踪结果。

图3为3种算法在6个视频序列上的跟踪效果图, 从实验结果可以看出, 本文提出的基于全局背

4.2 定量分析

为定量地评价跟踪算法的性能, 本文采用Precision Plot of OPE(精度曲线图)和Success Plots of OPE(成功率曲线图)对实验结果进行分析。精度曲线图纵轴表示距离精度, 横轴表示中心误差, 选取阈值为20像素。成功率曲线图纵轴表示成功率, 横轴表示重叠率阈值, 选取阈值为0.5。

(1)整体性能分析: 使用51个标准视频序列对5种跟踪算法进行整体性能分析, 并计算所有视频序列的平均精度和成功率, 如图1所示。

从图1中可以看出, 本文算法在总体的距离精度和成功重叠率两个指标上都取得第1名。本文算

表2 视频序列及其描述

序列	帧数	场景特征
Skiing	81	形变、快速运动、复杂背景
Football	81	运动模糊、背景干扰、形变
Freeman	297	背景干扰、形变
Singer	351	背景干扰、光照变化、尺度变化
Jumping	313	运动模糊、快速运动
Deer	71	快速运动、运动模糊

景与特征降维的目标跟踪算法拥有以下几个优点:

(1)对于形变和复杂背景干扰, 具有较好的抵抗能力。如Skiing和Freeman序列所示, 当目标出现形变和复杂背景干扰时, Staple算法和KCF算法均出现丢失或明显偏差情况; 而本文算法充分利用目标周围的背景信息建模, 能有效突出前景目标, 抑制背景干扰, 从而取得了最优跟踪结果。

(2)对于快速运动和运动模糊具有较强的适应能力。从结果来看, 本文通过自适应特征融合能更好地适应目标表观变化, 达到鲁棒跟踪的效果。

5 总结

针对相关滤波算法存在的问题, 本文提出了一种基于全局背景与特征降维的视觉跟踪算法。该算

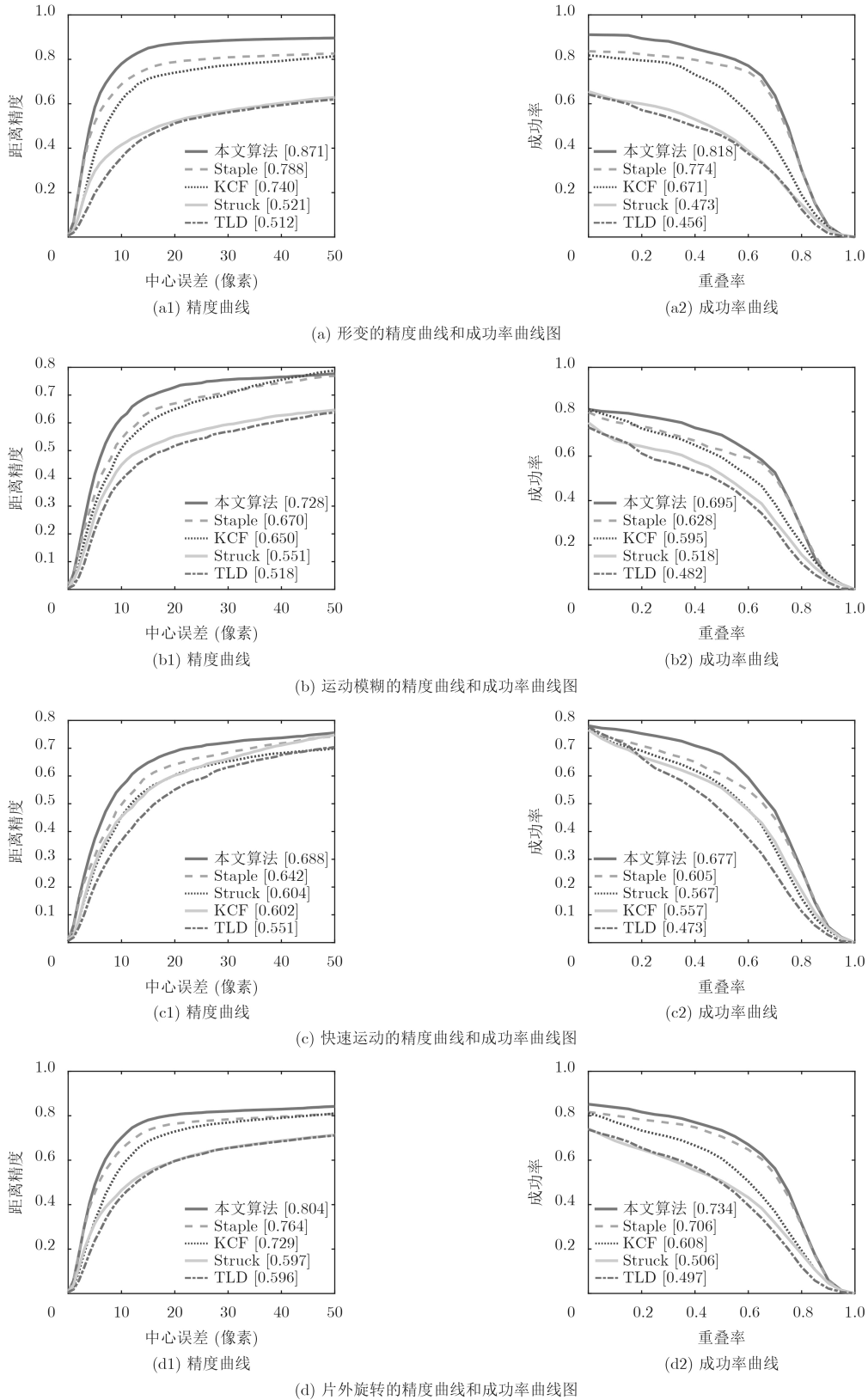


图 2 各属性的精度曲线和成功率曲线图

法充分利用目标周围的背景信息，使得训练的分类器更关注于目标与背景的差异性，从而避免了形变情况下背景中相似物的干扰；然后利用降维特征对

分类器模型进行参数更新，能有效减少运算复杂度；最后在响应图层面将HOG特征与颜色特征自适应融合，以确定最终的目标位置，解决了相关滤



图3 跟踪效果对比图

波算法对快速运动敏感的问题, 保证了快速运动和运动模糊情况下可靠的跟踪。实验结果表明: 本文算法能较好处理形变、运动模糊和快速运动等问题。

参 考 文 献

- [1] 齐苑辰, 吴成东, 陈东岳, 等. 基于稀疏表达的超像素跟踪算法[J]. 电子与信息学报, 2015, 37(3): 529–535. doi: [10.11999/JEIT140374](https://doi.org/10.11999/JEIT140374).
 QI Yuanchen, WU Chengdong, CHEN Dongyue, *et al.* Superpixel tracking based on sparse representation[J]. *Journal of Electronics & Information Technology*, 2015, 37(3): 529–535. doi: [10.11999/JEIT140374](https://doi.org/10.11999/JEIT140374).
- [2] JI Hui, LING Haibin, WU Yi, *et al.* Real time robust L1 tracker using accelerated proximal gradient approach[C]. IEEE International Conference on Computer Vision and Pattern Recognition, Providence, USA, 2012: 1830–1837.
- [3] KWON J and LEE K M. Tracking by sampling trackers[C]. International Conference on Computer Vision, Barcelona, Spain, 2011: 1195–1202.
- [4] LIU Baiyang, HUANG Junzhou, KULIKOWAKI C, *et al.* Robust visual tracking using local sparse appearance model and k-selection[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, 35(12): 2968–2981. doi: [10.1109/tpami.2012.215](https://doi.org/10.1109/tpami.2012.215).
- [5] HARE S, SAFFARI A, and TORR P H S. Struck: Structured output tracking with kernels[C]. International Conference on Computer Vision, Barcelona, Spain, 2011: 263–270.
- [6] KALAL Z, MATAS J, and MIKOLAJCZYK K. P-N learning: Bootstrapping binary classifiers by structural constraints[C]. IEEE International Conference on Computer Vision and Pattern Recognition, San Francisco, USA, 2010: 49–56.
- [7] BABENKO B, YANG Minghsuang, and BELONGIE S. Visual tracking with online multiple instance learning[C]. IEEE International Conference on Computer Vision and Pattern Recognition, Miami, USA, 2009: 983–990.
- [8] 孙航, 李晶, 杜博, 等. 基于多阶段学习的相关滤波目标跟踪[J]. 电子学报, 2017, 45(10): 2337–2342. doi: [10.3969/j.issn.0372-2112.2017.10.004](https://doi.org/10.3969/j.issn.0372-2112.2017.10.004).
 SUN Hang, LI Jing, DU Bo, *et al.* Correlation filtering target tracking based on online multi-lifespan learning[J]. *Acta Electronica Sinica*, 2017, 45(10): 2337–2342. doi: [10.3969/j.issn.0372-2112.2017.10.004](https://doi.org/10.3969/j.issn.0372-2112.2017.10.004).
- [9] BOLME D S, BEVERIDGE J R, DRAPER B A, *et al.* Visual object tracking using adaptive correlation filters[C]. IEEE International Conference on Computer Vision and Pattern Recognition, San Francisco, USA, 2010: 2544–2550.
- [10] HENRIQUES J F, CASEIRO R, MARTINS P, *et al.* Exploiting the circulant structure of tracking-by-detection with kernels[C]. Proceedings of European Conference on Computer Vision, Florence, Italy, 2012: 702–715.
- [11] HENRIQUES J F, RUI C, MARTINS P, *et al.* High-speed tracking with kernelized correlation filters[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2015, 37(3): 583–596. doi: [10.1109/tpami.2014.2345390](https://doi.org/10.1109/tpami.2014.2345390).
- [12] DANELLJAN M, HÄGER G, KHAN F, *et al.* Accurate scale estimation for robust visual tracking[C]. Proceedings of British Machine Vision Conference, Nottingham, UK, 2014: 65.1–65.11.
- [13] MUELLER M, SMITH N, and GHANEM B. Context-aware correlation filter tracking[C]. IEEE International Conference

- on Conference on Computer Vision and Pattern Recognition, Honolulu, USA, 2017: 1387–1395.
- [14] BERTINETTO L, VALMADRE J, GOLODETZ S, *et al.* Staple: Complementary learners for real-time tracking[C]. IEEE International Conference on Conference on Computer Vision and Pattern Recognition, Las Vegas, USA, 2016: 1401–1409.
- [15] 邬战军, 牛敏, 许冰, 等. 基于谱回归特征降维与后向传播神经网络的识别方法研究[J]. 电子与信息学报, 2016, 38(4): 978–984. doi: [10.11999/JEIT150781](https://doi.org/10.11999/JEIT150781).
- WU Zhanjun, NIU Min, XU Bing, *et al.* Research on recognition method based on spectral regression and back propagation neural network[J]. *Journal of Electronics & Information Technology*, 2016, 38(4): 978–984. doi: [10.11999/JEIT150781](https://doi.org/10.11999/JEIT150781).
- [16] 侯志强, 张浪, 余旺盛, 等. 基于快速傅里叶变换的局部分块视觉跟踪算法[J]. 电子与信息学报, 2015, 37(10): 2397–2404. doi: [10.11999/JEIT150183](https://doi.org/10.11999/JEIT150183).
- HOU Zhiqiang, ZHANG Lang, YU Wangsheng, *et al.* Local patch tracking algorithm based on fast fourier transform[J]. *Journal of Electronics & Information Technology*, 2015, 37(10): 2397–2404. doi: [10.11999/JEIT150183](https://doi.org/10.11999/JEIT150183).
- [17] DANELLJAN M, HAGER G, KHAN F S, *et al.* Discriminative scale space tracking[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2017, 39(8): 1561–1575. doi: [10.1109/tpami.2016.2609928](https://doi.org/10.1109/tpami.2016.2609928).
- [18] 熊昌镇, 赵璐璐, 郭芬红. 自适应特征融合的核相关滤波跟踪算法[J]. 计算机辅助设计与图形学学报, 2017, 29(6): 1068–1074. doi: [10.3969/j.issn.1003-9775.2017.06.012](https://doi.org/10.3969/j.issn.1003-9775.2017.06.012).
- XIONG Changzhen, ZHAO Lulu, and GUO Fenhong. Kernelized correlation filters tracking based on adaptive feature fusion[J]. *Journal of Computer-Aided Design & Computer Graphics*, 2017, 29(6): 1068–1074. doi: [10.3969/j.issn.1003-9775.2017.06.012](https://doi.org/10.3969/j.issn.1003-9775.2017.06.012).
- [19] WU Yi, LIM J, and YANG Minghsuan. Online object tracking: a benchmark[C]. Proceedings of the Computer Vision and Pattern Recognition, Portland, USA, 2013: 2411–2418.
- 孙彦景: 男, 1977年生, 教授, 博士生导师, 研究方向为无线网络、视频目标跟踪、人工智能、信息物理系统.
- 王赛楠: 女, 1992年生, 硕士生, 研究方向为视频目标跟踪.
- 石韞开: 男, 1994年生, 硕士生, 研究方向为人工智能.
- 云霄: 女, 1986年生, 讲师, 研究方向为视频目标跟踪和人工智能.
- 施文娟: 女, 1981年生, 博士生, 研究方向为视频质量评价.