

基于特征融合与分类器在线学习的目标跟踪算法

胡秀华[†], 郭雷, 李晖晖

(西北工业大学 自动化学院, 西安 710129)

摘要: 为了解决目标在复杂环境下表观变化引起的跟踪漂移问题, 提出一种基于多特征融合与分类器在线学习的目标跟踪算法. 该算法利用不同表观特征训练子分类器, 通过构建损失函数求得各子分类器可置信度, 进而加权融合子预测结果, 得到当前帧最佳目标状态估计; 同时, 依据最近-最远边界原则和协同训练理论粗更新训练样本集, 并通过精选择准则得到更具代表性的训练样本集, 实现子分类器自适应更新. 实验结果表明, 所提出的算法在多种典型测试场景中都能取得较鲁棒的跟踪效果.

关键词: 目标跟踪; 特征融合; 可置信度; 在线学习

中图分类号: TP391

文献标志码: A

Object tracking algorithm based on feature fusion and classifier online learning

HU Xiu-hua[†], GUO Lei, LI Hui-hui

(College of Automation, Northwestern Polytechnical University, Xi'an 710129, China)

Abstract: To solve the tracking drift problem caused by object appearance change in complex environments, the paper proposes an object tracking algorithm on the basis of multi-feature fusion and classifier online learning. The algorithm trains the sub-classifier with different apparent features, and calculates the reliability of each classifier by building the loss function, and then the optimum target state estimation by means of the weighted fusion prediction results of each sub-classifier is obtained. Meanwhile, it updates the training sample set coarsely according to the nearest-farthest boundary principle as well as the co-training theory, and gets more representative ones with the refined selection criterion, which further updates the sub-classifier adaptively. Experimental evaluations demonstrate that the proposed algorithm achieves favorable tracking performance against state-of-the-art methods on various typical testing scenarios.

Keywords: object tracking; feature fusion; reliability; online learning

0 引言

目标跟踪作为计算机视觉领域的重要研究方向, 为高层次视觉分析与理解奠定了基础. 然而, 受背景、光照、旋转、形变、尺寸、遮挡及运动模糊等复杂因素影响, 目标表观动态变化, 研究能够自适应复杂环境的鲁棒跟踪算法具有重要意义^[1-2].

现有目标跟踪算法大多分为3个模块, 包括表观描述运动估计和模型更新. 文献[3]指出, 特征描述在跟踪中起着关键作用, 运动模型、模型更新策略和整体后处理过程也在很大程度上影响着跟踪效果; 文献[4]将目标跟踪看作多任务稀疏学习问题, 考虑了候选样本之间的相关性, 联合稀疏表述各候选样本, 但是采用的整体稀疏表示模板不能有效处理目标遮

挡; 文献[5]利用整体模板和局部稀疏表示构建联合模型, 并运用最新观测和初始模板完成了模板更新, 该算法综合性能良好, 但不能有效适应目标剧烈运动和形变, 且算法实时性差; 文献[6]通过在线更新支持向量集合, 给出了基于结构输出预测的自适应跟踪框架, 并指出利用组合多种特征的核函数可以改善跟踪性能, 但是该算法采用固定尺度的密集采样策略, 在目标存在尺度变化时容易将背景干扰信息更新进模板; 文献[7]利用当前时刻与部分过往时刻形成的跟踪器组建了专家组, 根据熵正则化优化函数选择判别性最强的专家预测值作为输出结果, 该算法可以有效地从复杂背景中提取目标信息, 但在处理形变和运动模糊时性能减弱; 文献[8]构造了基于分块多特征描

收稿日期: 2016-05-18; 修回日期: 2016-08-18.

基金项目: 国家自然科学基金项目(61273362); 国家自然科学基金重点项目(61333017).

作者简介: 胡秀华(1988—), 女, 博士生, 从事目标跟踪、多传感器信息融合的研究; 郭雷(1956—), 男, 教授, 博士生导师, 从事图像处理、模式识别与智能系统等研究.

[†]通讯作者. E-mail: huxh@mail.nwpu.edu.cn

述子的检测器和基于卡尔曼滤波的跟踪器,并利用深度直方图信息估计目标尺寸与遮挡情况;文献[9]设计了融合多种特征的贝叶斯框架,但是该算法假设各特征条件独立,且没有考虑不同观测模型之间的可靠性和相关性,不能充分利用多特征之间的互补性;文献[10]通过分解的观测模型和运动模型构建了基本跟踪器,并基于交互马尔科夫链蒙特卡洛框架完成并行运行跟踪器的信息关联,但是该算法没有考虑各跟踪器的可靠性;基于协同训练框架,文献[11]给出了一种结合生成式模型与判别式分类器的跟踪方法,但是该算法过于依赖目标表观稳定性,不能有效处理表观剧烈变化和部分遮挡问题.现有研究成果虽然在解决目标表观动态变化问题时考虑了表观特征描述和模板更新等改进方法,但是在多特征信息的充分利用以及模型自适应更新方面还有待深入研究.

针对目标在剧烈表观变化时容易出现跟踪漂移的问题,本文设计一种基于多特征融合与分类器在线学习的目标跟踪算法.首先,利用支持向量机原理训练得到基于不同表观特征子分类器;然后,依据最大后验估计理论,构建基于对数似然与条件熵的损失函数,求得各子分类器的可信度,并通过加权融合各子分类器预测结果得到当前帧最佳目标运动估计;最后,设计训练样本集与分类器更新准则,借助最近-最远边界原则与协同训练理论粗更新训练样本集,并通过精选择得到更具有代表性的训练样本集,进而更新得到强泛化能力的分类器,可以更有效地将目标从背景中分离出来.

1 问题描述

为充分利用不同特征之间的互补特性,选用灰度特征和方向梯度直方图(HOG)描述目标表观,其中灰度特征能够保留目标表观原始信息,HOG适合处理目标形变、运动模糊等表观变化.假设两种特征条件独立,则可以训练得到两个相对独立的子分类器.

目标跟踪问题可以视为目标和背景的二分类问题,对于单个特征描述子,基于支持向量机理论训练一个分类超平面,将目标从背景中分离出来^[12].

对于第*i*类特征描述子,目标训练样本特征向量集 \mathcal{X}_i 为 $\{\mathbf{x}_{i1}, \mathbf{x}_{i2}, \dots, \mathbf{x}_{i(P+Q)}\}$,训练样本标签集 \mathcal{Y}_i 为 $\{y_{i1}, y_{i2}, \dots, y_{i(P+Q)}\}$, \mathbf{x}_{ij} 为第*i*类特征描述子的第*j*个训练样本特征向量, $y_{ij} \in (-1, +1)$ 为相应的训练标签, $i = 1, 2, j = 1, 2, \dots, P + Q, P$ 为正样本数, Q 为负样本数.定义子分类决策函数为

$$f(\mathbf{x}_i) = \mathbf{w}_i^T \mathbf{x}_i + b_i, \quad (1)$$

其中 \mathbf{w}_i, b_i 为决定超平面的参数.

通过最大化分类间隔得出最小化目标函数

$$\min_{\mathbf{w}_i, b_i, \xi_i} \frac{1}{2} \|\mathbf{w}_i\|^2 + C_i \sum_{j=1}^{P+Q} \xi_{ij}. \quad (2)$$

其中: $y_{ij}(\mathbf{w}_i^T \mathbf{x}_{ij} + b_i) \geq 1 - \xi_{ij}, \xi_{ij} \geq 0, \xi_{ij}$ 为松弛变量;参数 C_i 为平衡系数.

通过求解对偶问题,式(2)可以等价表示为

$$\max_{\alpha_i} \sum_{j=1}^{P+Q} \alpha_{ij} - \frac{1}{2} \sum_{j,l=1}^{P+Q} \alpha_{ij} \alpha_{il} y_{ij} y_{il} \mathbf{x}_{ij}^T \mathbf{x}_{il}. \quad (3)$$

其中: $0 \leq \alpha_{ij} \leq C_i, \sum_{j=1}^{P+Q} \alpha_{ij} y_{ij} = 0, \alpha_{ij} \geq 0$ 为拉格朗日乘子.

利用序列最小优化算法求得拉格朗日乘子向量 $[\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{i(P+Q)}]^T$,进而导出子分类决策函数系数 \mathbf{w}_i^* 和 b_i^* .

对候选样本 $\hat{\mathbf{x}}_{ik}$ 进行分类时,把 $\hat{\mathbf{x}}_{ik}$ 代入 $f(\hat{\mathbf{x}}_{ik}) = \mathbf{w}_i^T \hat{\mathbf{x}}_{ik} + b_i$ 中,然后依据 $f(\hat{\mathbf{x}}_{ik})$ 值进行判别.因此,由 $\mathbf{w}_i^* = \sum_{j=1}^{P+Q} \alpha_{ij} y_{ij} \mathbf{x}_{ij}$ 及式(1)可得

$$f(\hat{\mathbf{x}}_{ik}) = \sum_{j=1}^{P+Q} \alpha_{ij} y_{ij} K(\mathbf{x}_{ij}, \hat{\mathbf{x}}_{ik}) + b_i. \quad (4)$$

其中:核函数 $K(\mathbf{x}_{ij}, \hat{\mathbf{x}}_{ik})$ 表征 $\mathbf{x}_{ij}, \hat{\mathbf{x}}_{ik}$ 在特征空间的内积, $\hat{\mathbf{x}}_{ik}$ 为基于第*i*类特征描述子的候选样本特征向量, $i = 1, 2, k = 1, 2, \dots, \Gamma, \Gamma$ 为候选样本数.本文中选取高斯核函数训练分类器,即有 $K(\mathbf{x}_{ij}, \hat{\mathbf{x}}_{ik}) = \exp(-\|\mathbf{x}_{ij} - \hat{\mathbf{x}}_{ik}\|^2 / 2\sigma^2)$, σ 为高斯函数方差.对于候选样本 $\hat{\mathbf{x}}_{ik}$ 的类别预测问题,只需要计算它与训练样本集的内积即可.

由多个子分类器组成的跟踪系统,目标预测位置取决于各分类器的共同作用结果.由于不同特征之间条件独立,可以通过加权各子分类器结果得到融合的目标位置,但需要深入研究如何自适应利用不同特征的优势,以及设计鲁棒的分类器在线学习策略.

2 算法的提出

针对上述问题,本文给出一种新算法.该算法综合利用互补特征进行目标表观描述,在利用支持向量机理论构建各子分类器的基础上,依据最大后验估计求得各子分类器的可信度,并通过加权融合各子分类器预测结果得到最佳候选目标,同时,设计分类器在线学习及相应更新策略,可以较准确地将目标从背景中分离出来.算法原理如图1所示.

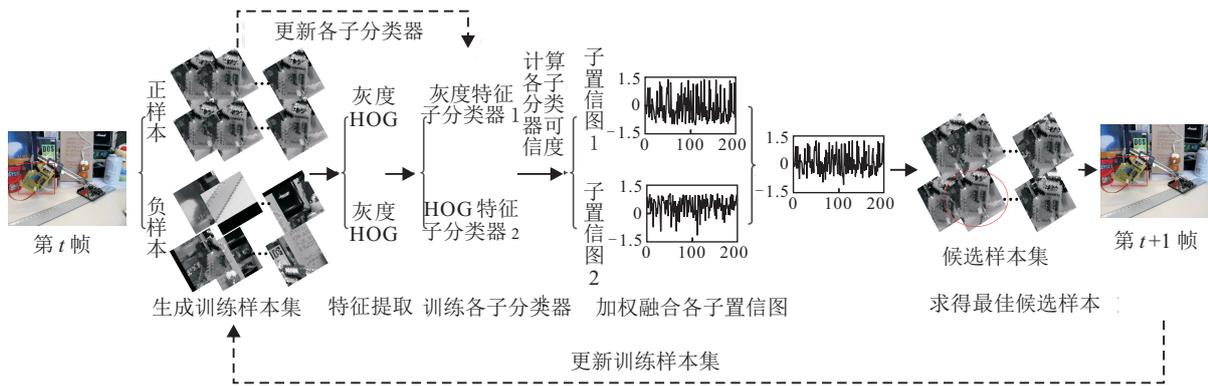


图1 算法原理

所提出的算法的整体步骤描述如下。

输入: 初始训练样本集 $\mathcal{X}_i, \mathcal{Y}_i, i = 1, 2$

For $t = 2$ to frame end

从训练样本中提取灰度与HOG特征

由式(1)~(4)训练得到第 $t - 1$ 帧各特征对应的子分类器 $f(x_i)$

构建测试样本集 $\tilde{\mathcal{X}}_i$, 依据式(8)~(13)求得各子分类器可信度 ω_i

采集 Γ 个候选样本, 由式(5)~(7)和(14)求得目标在第 t 帧的最佳状态估计值 \hat{s}_t

依据在线学习策略更新训练样本集 \mathcal{U}_{t-1} , 进而更新得到第 t 帧各子分类器 $f(x_i)$

End

输出: 目标最佳运动状态估计 $\{\hat{s}_t\}_{t=2,3,\dots}$

所提出算法的各部分具体描述将在下文给出。

2.1 目标表观描述

将图像转化为灰度图像, 手动标出初始矩形目标区域, 并归一化为 $m \times m$ 个像素。记目标区域中心为 $X(h, l)$, 在目标中心周围 $\|X_p - X(h, l)\| \leq r_o$ 圆形区域范围内, 密集采样提取 P 个正样本, 记作正样本集 \mathcal{N}_P 。其中: X_p 为第 p 个正样本的中心位置, r_o 为圆形区域半径, $\mathcal{N}_P = \{N_1, N_2, \dots, N_P\}$, $N_p \in \mathbf{R}^{m \times m}$ 为第 p 个正样本, $p = 1, 2, \dots, P$ 。在目标周围 $r_1 \leq \|X_q - X(h, l)\| \leq r_2$ 环形区域范围内, 采样提取 Q 个负样本, 记作负样本集 \mathcal{N}_Q 。其中: X_q 为第 q 个负样本的中心位置; r_1, r_2 为环形区域的内、外半径, $r_o < r_1 < r_2$; $\mathcal{N}_Q = \{N_1, N_2, \dots, N_Q\}$, $N_q \in \mathbf{R}^{m \times m}$ 为第 q 个负样本, $q = 1, 2, \dots, Q$ 。

归一化目标样本大小为 32×32 个像素, 提取灰度特征时, 通过滑动采样得到各图像块, 每个图像块为 8×8 个像素, 滑动步长为 4 个像素, 得到大小为 3136 维的灰度特征向量。提取 HOG 特征描述子时, 每个细胞单元为 4×4 个像素, 每个区间块大小为 8×8 个像素, 包含 4 个细胞单元, 块的滑动采样步长为 4 个像素,

采用 9 个直方图通道描述每个细胞单元, 并归一化每个区间块的梯度直方图, 最后每个样本得到大小为 1764 维的 HOG 特征向量。

2.2 目标状态估计

在第 t 帧, 令 s_t 表示目标运动状态变量, 给定观测向量 $\mathcal{O}_t = [o_1, o_2, \dots, o_t]$, 基于贝叶斯理论框架, 递归地估计隐状态变量 s_t 的值^[13], 即

$$p(s_t | \mathcal{O}_t) \propto p(o_t | s_t) \cdot \int p(s_t | s_{t-1}) p(s_{t-1} | \mathcal{O}_{t-1}) ds_{t-1}. \quad (5)$$

其中: $p(s_t | s_{t-1})$ 为目标在连续帧之间的运动模型; $p(o_t | s_t)$ 为在状态 s_t 下对观测量 o_t 进行估计的观测模型, $t = 2, 3, \dots$ 。

目标在第 t 帧的最佳运动状态可以通过最大化后验概率估计获得, 即目标运动状态可表示为

$$\hat{s}_t = \arg \max_{s_t^\tau} p(s_t^\tau | \mathcal{O}_t). \quad (6)$$

其中: s_t^τ 表示第 τ 个候选样本的状态变量, $\tau = 1, 2, \dots, \Gamma$, Γ 为采样粒子数。

定义目标状态变量 $s_t = [x_t, y_t, \theta_t, s_t, \eta_t, \psi_t]$, 其中 $x_t, y_t, \theta_t, s_t, \eta_t, \psi_t$ 分别表示目标在水平与竖直方向上的平移量、旋转角、尺度变化量、高宽比和斜切角, 6 个仿射参数相互独立, 采用随机游走模型描述连续两帧之间目标状态的转换, 即

$$p(s_t | s_{t-1}) = N(s_t; s_{t-1}, \Sigma). \quad (7)$$

其中: $\Sigma = \text{diag}\{\sigma_x^2, \sigma_y^2, \sigma_\theta^2, \sigma_s^2, \sigma_\eta^2, \sigma_\psi^2\}$, 对角线上元素依次为各仿射参数的方差。

在第 t 帧, 依据第 $t - 1$ 帧的目标跟踪结果, 利用粒子滤波采样得到 Γ 个候选样本, 归一化候选样本 y_τ 为 $m \times m$ 像素并进行特征提取, 得到灰度特征向量和 HOG 特征向量, $\tau = 1, 2, \dots, \Gamma$ 。

为了从运动模型得到的大量候选样本中找到要跟踪的目标, 利用训练得到的两个子分类器分别对各候选样本进行类别判定, 得到表征目标位置似然函数的子置信图 $c(\hat{x}_i)$ 。由于各子分类器对于目标表观变

化的鲁棒程度有所差别,需要衡量不同时刻各子分类器对候选样本类别判定的可靠程度.

为判别各子分类器的可信度,依据第1帧和最新 ζ 帧中得分最高的样本及其周围区域构建测试集^[7], Π 为测试样本数,记第 i 类特征描述子对应的测试样本特征向量集 $\tilde{\mathcal{X}}_i$ 为 $\{\tilde{\mathbf{x}}_{i1}, \tilde{\mathbf{x}}_{i2}, \dots, \tilde{\mathbf{x}}_{i\Pi}\}$,由分类器求得的相应标签集 $\tilde{\mathcal{Y}}_i$ 为 $\{\tilde{y}_{i1}, \tilde{y}_{i2}, \dots, \tilde{y}_{i\Pi}\}$,对应的基准标签集 \mathcal{Z}_i 为 $\{z_{i1}, z_{i2}, \dots, z_{i\Pi}\}$.

定义损失函数 Ψ_i ,通过引入熵正则化项,有

$$\Psi_i(\tilde{\mathcal{X}}_i, \tilde{\mathcal{Y}}_i) = -L(\theta_i; \tilde{\mathcal{X}}_i, \tilde{\mathcal{Y}}_i) + \lambda H(\mathcal{Z}_i | \tilde{\mathcal{X}}_i, \tilde{\mathcal{Y}}_i; \theta_i). \quad (8)$$

其中: θ_i 为模型参数, $L(\theta_i; \tilde{\mathcal{X}}_i, \tilde{\mathcal{Y}}_i)$ 为 θ_i 的对数似然函数, $H(\mathcal{Z}_i | \tilde{\mathcal{X}}_i, \tilde{\mathcal{Y}}_i; \theta_i)$ 为测试样本标签真值关于训练样本和可能标签集的条件熵, λ 为正则化参数.

基于最大后验估计理论,构建似然函数与条件熵,即

$$L(\theta_i; \tilde{\mathcal{X}}_i, \tilde{\mathcal{Y}}_i) = \max_{\tilde{\mathcal{Y}} \subset \mathcal{Z}_i} \log P(z_i | \tilde{\mathbf{x}}_i; \theta_i), \quad (9)$$

$$H(\mathcal{Z}_i | \tilde{\mathcal{X}}_i, \tilde{\mathcal{Y}}_i; \theta_i) = -\sum_{\eta=1}^{\Pi} P(z_{i\eta} | \tilde{\mathbf{x}}_{i\eta}, \tilde{y}_{i\eta}; \theta_i) \cdot \log P(z_{i\eta} | \tilde{\mathbf{x}}_{i\eta}, \tilde{y}_{i\eta}; \theta_i), \quad (10)$$

$$P(z_i | \tilde{\mathbf{x}}_i; \theta_i) = \prod_{\eta=1}^{\Pi} P(z_{i\eta} | \tilde{\mathbf{x}}_{i\eta}; \theta_i), \quad (11)$$

$$P(z_{i\eta} | \tilde{\mathbf{x}}_{i\eta}, \tilde{y}_{i\eta}; \theta_i) = \frac{\delta_{\tilde{y}_{i\eta}}(z_{i\eta}) P(z_{i\eta} | \tilde{\mathbf{x}}_{i\eta}; \theta_i)}{\sum_{z'_{i\eta} \in \mathcal{Z}_i} \delta_{\tilde{y}_{i\eta}}(z'_{i\eta}) P(z'_{i\eta} | \tilde{\mathbf{x}}_{i\eta}; \theta_i)}. \quad (12)$$

设第 i 个子分类器的分类误差为 ε_i ,则相应的可信度表示为

$$\omega_i = \frac{\varepsilon_{3-i} + \varsigma}{\varepsilon_i + \varepsilon_{3-i} + \varsigma}. \quad (13)$$

其中: $\varepsilon_i = \kappa \Psi_i(\tilde{\mathcal{X}}_i, \tilde{\mathcal{Y}}_i)$, κ 表示各子分类器的分类误差与损失函数的正相关度, $i = 1, 2$, ς 为较小的常量,用以避免各分类器损失均为零.

利用得到的各子分类器可信度,加权融合相应的子置信图,总置信图中的峰值即为目标最可能出现的位置,则在第 t 帧,目标的观测模型可表示为

$$p(\mathbf{o}_t | \mathbf{s}_t) = \max_{\Gamma} \sum_{i=1}^2 \omega_i c(\hat{\mathbf{x}}_{i\tau}). \quad (14)$$

其中: $c(\hat{\mathbf{x}}_{i\tau})$ 为置信图中第 i 个子分类器得到的关于第 τ 个候选样本的值, $\tau = 1, 2, \dots, \Gamma$.

2.3 分类器在线学习策略

为适应目标表观动态变化,训练样本集与各子分类器需要适时更新,并且训练样本的代表性和分类器的更新策略直接影响分类器的性能.传统的基于各

子分类器预测结果进行自训练的更新方式容易积累误差,基于主动学习和协同训练理论思想,选择信息含量丰富和表观能力强的样本加入训练样本集,同时将不同特征表述下选取的优胜样本交互使用,可以提升各子分类器的鲁棒性能.

越靠近超平面的候选样本不确定性越高,否则反之,因此依据最近-最远边界原则对新增训练样本进行粗选择.对于第 i 个子分类器,在第 t 帧,由 $t-1$ 帧得到的子分类器对各候选样本进行类别标记,选取得分最低的 $R-2$ 个负候选样本作为新增负训练样本,同时选取得分最高的正、负候选样本分别作为新增正、负训练样本,得到初始新增训练样本集 \mathcal{R}_i ,有效地保证了新增训练样本的代表性.

设定总训练样本数 $\hat{\Omega}$ 不超过 $P+Q+2R$ 个,记第 $t-1$ 帧时总训练样本集为 \mathcal{U}_{t-1} ,在第 t 帧,将各子分类器选取的新训练样本与原训练样本合并,为避免引入背景干扰信息,同时使更新的分类器适应剧烈表观变化,从4个方面对粗总样本集 $\{\mathcal{U}_{t-1}, \mathcal{R}_1, \mathcal{R}_2\}$ 进行精选择,具体描述如下.

1) 为避免异常值的影响,对训练样本 r 进行判定,若 $\max(0, 1 - y_{ir} f(\hat{\mathbf{x}}_{ir})) > 0$,则说明样本 r 异常,需要从粗选择的总训练样本集中优先去除样本 r ;

2) 因背景信息变化剧烈,优先移除粗总训练样本中对分类器贡献小的最早生成的部分负样本;

3) 总训练样本集中的正样本数相对较少且信息单一,当正样本总数超过设定阈值 Δ_1 时移除可信度最低的子分类器选取的部分正样本;

4) 支持向量集对分类器的泛化能力影响重大,当支持向量总数超过设定上限 Δ_2 时,加权合并同类别的最近邻样本.

随着各子分类器的在线学习及适时更新,其相应的权值也发生变化,得到更加准确的融合置信图,进而有效提高目标跟踪的鲁棒性.

2.4 时间复杂度分析

在计算时间复杂度方面,所提出的算法耗时主要集中在各子分类器可信度计算、最佳候选样本状态估计、训练样本集与分类器在线更新部分.各子分类器可信度计算耗时主要取决于式(8),计算复杂度可表示为 $O(\Pi \log \Pi)$, Π 为测试样本个数;候选样本集中最佳目标运动状态估计的计算耗时主要在于式(6),计算复杂度可表示为 $O(\Gamma \hat{\Delta}_{2,t-1} D_i)$, Γ 为候选样本总数, $\hat{\Delta}_{2,t-1}$ 为当前帧用于判定各候选样本类别的支持向量总数, D_i 为训练样本集中各样本的第 i 类子特征向量维数;训练样本集与各子分类器更新的计算耗时方面,主要取决于新增训练样本的粗选择、精

选择以及各子分类器的更新,当前帧计算复杂度可表示为 $O(RI + \Omega_t \Delta_{1,t} \Delta_{2,t} + \hat{\Omega}_t \hat{\Delta}_{2,t} D_i)$, R 为粗训练样本集中新增样本个数, Ω_t 为粗训练样本总数, $\Delta_{1,t}$ 为粗训练样本集中正样本总数, $\Delta_{2,t}$ 为粗训练样本集中支持向量总数, $\hat{\Omega}_t$ 为精选择得到的训练样本总数, $\hat{\Delta}_{2,t}$ 为当前帧支持向量总数.

3 实验结果与分析

实验中算法的实现环境为 Matlab 2012b, 计算机处理器为 i3-2130, 主频为 3.40 GHz, 内存为 4 GB. 设定正样本数 P 为 10, 负样本数 Q 为 200, 高斯核函数标准差 σ 为 0.2, ζ 为 5. 为了验证本文算法的有效性, 选择基准库中的 10 组具有自身形态变化、尺度变化、运动模糊、外界光照变化、背景干扰及遮挡等复杂情况的视频序列进行测试, 并与 9 种典型算法对比, 包括 VTD^[10]、LSK^[14]、MIL^[15]、MTT^[4]、SCM^[5]、Struck^[6]、STC^[16]、DSSmap^[17]、MEEM^[7] 算法, 并采用中心位置误差和边界框重叠率定量评价各跟踪算法性能. 图 2~图 11 给出了不同时刻不同算法在各测试序列中的中心位置误差值, 表 1 和表 2 给出了不同算法在各测试序列中的中心位置误差平均值和边界框重叠率平均值.

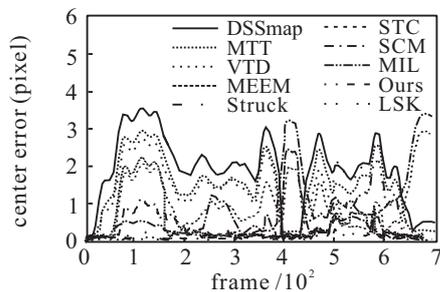


图 2 board

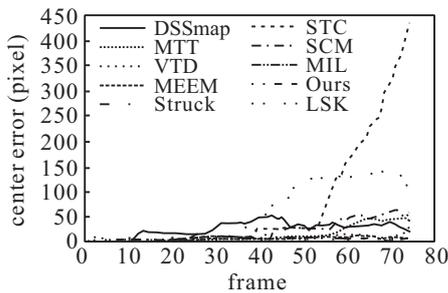


图 3 football 1

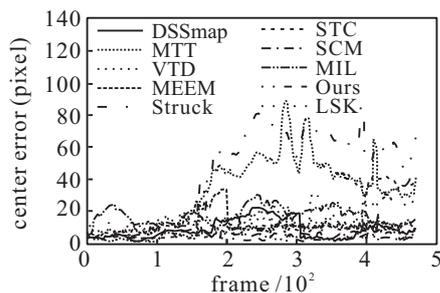


图 4 david

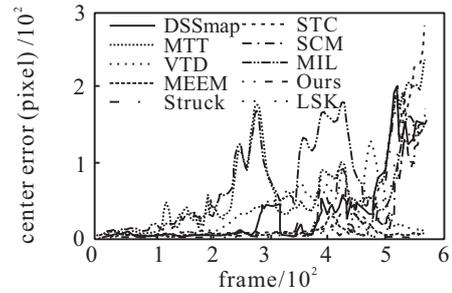


图 5 trellis

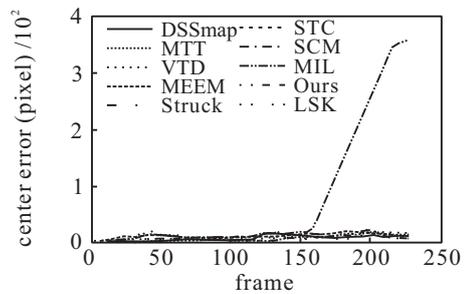


图 6 mountainBike

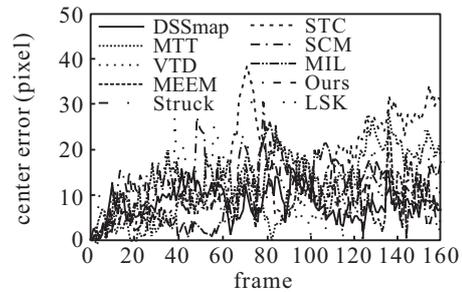


图 7 skater

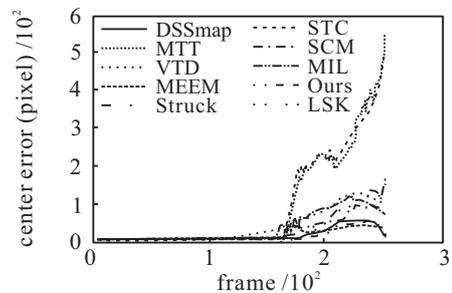


图 8 carScale

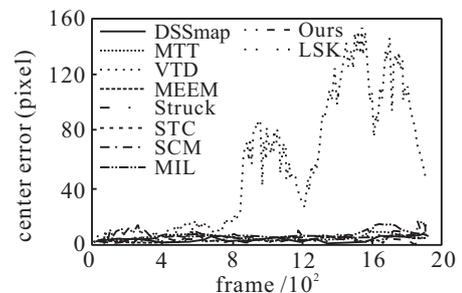


图 9 redTeam

由图 2~图 11 可以看出, 本文算法充分利用了不同特征之间的互补特性, 并通过所设计的在线学习策略得到强泛化能力的分类器, 保证了新算法在多种复杂环境下能有效实现目标的稳定跟踪, 适用性好.

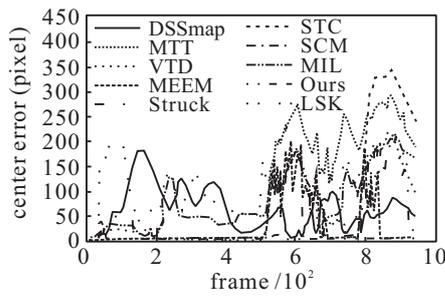


图 10 suv

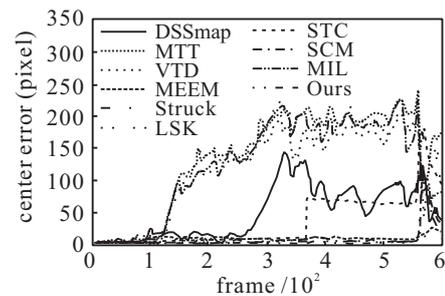


图 11 woman

表 1 中心位置误差平均值(pixel)

Sequence	VTD	LSK	MIL	MTT	SCM	Struck	STC	DSSmap	MEEM	Ours
board	125.174 6	50.7945	91.623 1	148.878 5	44.292 5	33.576 7	100.394 0	191.311 7	44.449 0	9.040 2
football1	7.436 5	58.220 4	5.617 8	13.539 0	20.019 5	5.465 0	72.573 8	26.146 5	5.061 0	5.578 6
david	11.586 0	12.022 8	16.859 8	31.684 3	4.636 8	42.800 8	12.163 6	8.401 2	11.166 6	3.619 3
trellis	32.250 0	20.307 2	71.468 6	64.709 8	22.650 8	25.638 2	33.753 9	33.241 5	6.148 8	2.811 2
mountainBike	9.784 4	10.352 8	73.023 0	7.830 2	10.451 1	12.600 0	7.032 4	9.062 0	12.981 2	6.948 4
skater	9.719 8	11.994 5	10.201 9	11.757 8	12.334 8	8.801 5	15.495 5	7.850 7	11.838 0	7.698 7
carScale	38.461 4	25.676 6	33.470 9	87.711 4	23.854 5	29.362 4	89.351 5	13.212 5	15.371 9	11.288 0
redTeam	52.935 2	3.099 5	6.651 2	4.259 9	3.672 7	4.357 2	3.446 0	3.101 3	3.534 7	2.823 2
suv	57.246 6	113.887 6	82.228 9	100.262 1	7.122 9	45.864 1	51.494 6	66.665 6	29.843 9	4.559 9
woman	118.862 7	131.053 7	125.293 3	136.671 2	7.762 3	4.167 6	32.467 1	50.728 3	9.450 2	6.851 5

表 2 边界框重叠率平均值

Sequence	VTD	LSK	MIL	MTT	SCM	Struck	STC	DSSmap	MEEM	Ours
board	0.266 7	0.617 4	0.406 7	0.206 3	0.576 3	0.652 8	0.315 8	0.144 7	0.611 0	0.731 1
football1	0.569 5	0.301 8	0.659 6	0.498 9	0.396 4	0.670 8	0.341 5	0.247 2	0.656 5	0.658 8
david	0.558 4	0.559 1	0.428 2	0.301 8	0.721 7	0.239 0	0.529 2	0.591 3	0.529 1	0.761 4
trellis	0.453 0	0.634 6	0.247 9	0.246 2	0.578 3	0.505 8	0.464 7	0.445 7	0.617 5	0.829 0
mountainBike	0.697 9	0.679 6	0.452 0	0.735 2	0.673 0	0.625 1	0.581 8	0.652 0	0.592 2	0.739 6
skater	0.541 6	0.575 9	0.612 7	0.545 7	0.592 9	0.633 3	0.409 1	0.631 4	0.573 9	0.646 9
carScale	0.429 8	0.578 5	0.409 7	0.485 3	0.620 4	0.431 3	0.420 2	0.721 3	0.405 2	0.735 1
redTeam	0.204 2	0.619 9	0.399 4	0.474 2	0.481 6	0.493 8	0.250 4	0.527 8	0.496 3	0.619 7
suv	0.452 7	0.050 3	0.202 8	0.468 6	0.714 2	0.536 9	0.513 8	0.175 7	0.612 3	0.773 6
woman	0.144 2	0.151 8	0.156 9	0.155 0	0.662 7	0.711 8	0.338 3	0.327 8	0.503 1	0.693 3

表1和表2中加粗数值代表最优结果,可以看出,本文算法在典型测试环境下取得了鲁棒跟踪效果,平均跟踪误差控制在12 pixels以内,平均边界框重叠率高于0.60,明显优于MEEM等算法。

各算法在各序列上的部分跟踪效果如图12所示。在图12中:黑色实线、黑色虚线、黑色点线分别表示VTD、LSK、MIL,白色实线、白色虚线、白色点线分别表示MTT、SCM、Struck,灰色实线、灰色虚线、灰色点线、灰色点实线分别表示STC、DSSmap、MEEM、Ours。

背景干扰:对于board序列,目标在背景混杂环境中摇动,伴随旋转形变和运动模糊,MTT等算法在跟踪过程中出现明显漂移乃至失败。在第488帧与581帧之间,目标反转,本文算法跟踪误差相对增大,但相比Struck和SCM等算法仍取得了较优的跟踪性能。对于football1序列,目标受混淆干扰和运动模糊影响严重,在第40帧目标出现运动模糊,只是利用了

局部信息的LSK不能有效适应背景干扰。第56帧目标突然向下,STC跟踪误差随之增大,Struck、MEEM、MIL和本文算法可以完整跟踪目标,其中,MEEM取得了较小的跟踪误差,Struck取得了较高的边界框重叠率。

光照变化:david序列存在明显光照变化,在第168帧,受运动模糊和旋转影响,本文算法跟踪误差短时增大,MTT和Struck出现明显跟踪漂移。整体上看,SCM和本文算法能有效实现完整跟踪,且本文算法取得了相对较优的跟踪性能。对于trellis序列,在第221帧,目标进入高光照区域,MIL和MTT跟踪误差增大。在第512帧之后,受光照变化和背景干扰影响,LSK和SCM等算法均出现不同程度的跟踪漂移,而本文算法采用的互补特征信息与在线学习策略可以充分利用目标和背景信息,取得了较为鲁棒的跟踪效果。

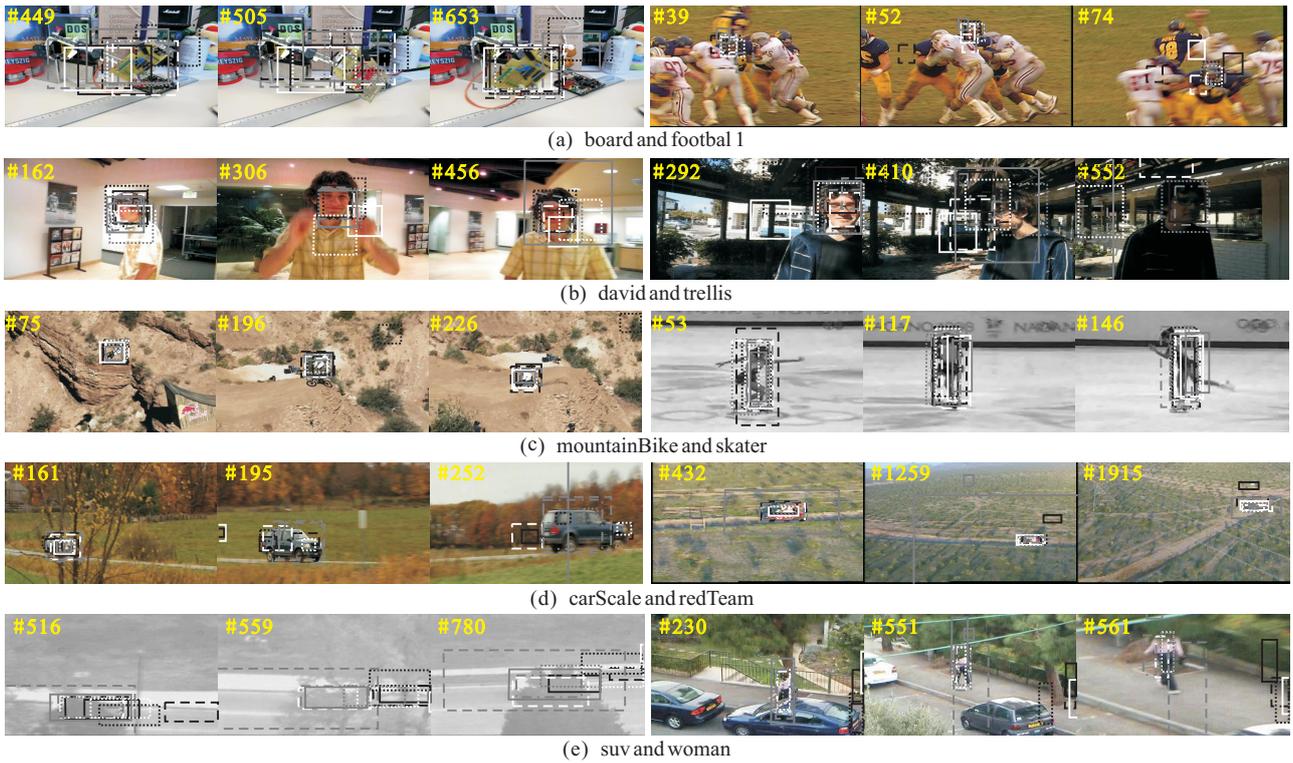


图 12 部分时刻不同算法在各测试视频序列上的跟踪效果

旋转变形: mountainBike 序列受旋转变形和背景干扰等影响, 在第 41 帧, 目标跳起, 多数算法跟踪误差增大. 在第 155 帧, 目标降落, MIL 出现明显跟踪漂移, 而利用多特征信息的本文算法和基于稀疏混合正则项表述的 MTT 算法跟踪误差相对较小, 较其他算法有明显优势. skater 序列中目标持续旋转变形, 在第 64~78 帧, 目标重心位置突变, 所有算法跟踪误差明显增加, 其中 STC 跟踪误差相对较大. 在第 84、85 以及 185 帧之后, 本文算法在准确估计目标位置上有所偏差, 但在其余时段跟踪精度相对较高, 整体上看, Struck、DSSmap 和本文算法平均跟踪性能良好.

尺度变化: carScale 序列中, 目标快速运动且尺寸逐步增大, 在第 157~178 帧, 受外界遮挡影响, STC 和 MTT 算法出现明显跟踪漂移. 在第 203 帧之后, 受尺寸增大和平面外旋转影响, 所有算法跟踪误差增大, 但 DSSmap 和本文算法误差增加相对较小. redTeam 序列中, 目标尺寸不断变化, 在第 803 帧, 目标尺寸相对较小, 分辨率较低, 基于整体模板且没有考虑各子跟踪器关联性的 VTD 出现明显跟踪漂移乃至失败, LSK 平均边界框重叠率相对较高, 但在第 1862 帧之后跟踪误差相对较大, 而 DSSmap、MEEM 和本文算法整体跟踪性能较优.

严重遮挡: suv 序列受外部遮挡干扰严重, 在第 29 帧, 目标逐渐淡出视野, LSK 算法出现跟踪失败. 从第 65 帧开始, DSSmap 跟踪误差增大乃至丢失目标, 直至第 402 帧目标重新进入扩大的搜索区域, 跟踪性能波动明显. 在第 511 帧, 受严重遮挡影响, MTT、Struck 和 MEEM 等算法也出现明显跟踪偏移, 而 SCM 和本文算法对目标运动趋势判定基本正确. woman 序列存在严重遮挡和明显形变, 在第 115 帧之后, VTD、LSK、MIL 和 MTT 算法均出现明显跟踪漂移. 在第 561 帧, 相机出现明显抖动, 目标尺寸增大并伴随模糊, 多数算法跟踪误差瞬间增加, 而 Struck 和 MEEM 跟踪性能相对稳定.

在算法运算速度方面, 选择适当数目的测试样本和支持向量数阈值可以很好地平衡消耗时间与性能提升. 本文算法中测试样本数 I 为 102, 候选样本数 Γ 为 200, 单个子分类器新增训练样本数 R 为 20, 正样本总数阈值 Δ_1 为 20, 支持向量总数上限 Δ_2 为 64, 表 3 给出了各算法在测试序列中的每帧平均耗时.

由定性和定量结果可知, 本文算法中给出的多特征融合方法与分类器在线学习策略虽然在一定程度上影响了实时性, 但却有效提高了算法的跟踪性能.

表 3 每帧平均耗时

算法	VTD	LSK	MIL	MTT	SCM	Struck	STC	DSSmap	MEEM	Ours
每帧平均耗时	0.1850	0.1865	0.0281	1.0279	1.8574	0.0519	0.0029	0.2683	0.1036	0.1467

4 结 论

本文充分利用不同特征之间的互补特性,设计了一种基于特征融合和分类器在线学习的目标跟踪算法.该算法利用支持向量机理论训练得到基于不同表现特征的子分类器,依据最大后验估计理论构建损失函数,求得各子分类器的可信度,通过加权融合各子分类器预测结果得到最佳目标运动状态估计;此外,给出了训练样本集和分类器更新策略,依据粗更新训练样本集及精选择准则,得到了更具代表性的训练样本集和更强泛化能力的分类器.实验结果和分析表明,所提出的算法能够有效解决目标在多种复杂环境下跟踪性能差的问题.

参考文献(References)

- [1] Wu Y, Lim J, Yang M H. Object tracking benchmark[J]. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 2015, 37(9): 1834-1848.
- [2] Kristan M, Matas J, Leonardis A, et al. The visual object tracking vot2015 challenge results[C]. *Proc of IEEE Conf on Computer Vision*. Santiago: IEEE Press, 2015: 564-586.
- [3] Wang N Y, Shi J P, Yeung D Y, et al. Understanding and diagnosing visual tracking systems[C]. *Proc of IEEE Conf on Computer Vision*. Santiago: IEEE Press, 2015: 3101-3109.
- [4] Zhang T, Ghanem B, Liu S, et al. Robust visual tracking via multi-task sparse learning[C]. *Proc of IEEE Conf on Computer Vision and Pattern Recognition*. Providence: IEEE Press, 2012: 2042-2049.
- [5] Zhong W, Lu H, Yang M H. Robust object tracking via sparsity-based collaborative model[C]. *Proc of IEEE Conf on Computer Vision and Pattern Recognition*. Providence: IEEE Press, 2012: 1838-1845.
- [6] Hare S, Saffari A, Torr P H S. Struck: Structured output tracking with kernels[C]. *Proc of IEEE Conf on Computer Vision*. Barcelona: IEEE Press, 2011: 263-270.
- [7] Zhang J M, Ma S G, Stan S. MEEM: Robust tracking via multiple experts using entropy minimization[C]. *Proc of the 13th European Conf on Computer Vision: Part VI*. Zurich: Springer International Publishing, 2014: 188-203.
- [8] 王丽佳, 贾松敏, 李秀智, 等. 分块多特征目标描述子的移动机器人目标跟踪[J]. *控制与决策*, 2016, 31(2): 337-342.
(Wang L J, Jia S M, Li X Z, et al. Person tracking for robot using patches-based-multi-cues representation[J]. *Control and Decision*, 2016, 31(2): 337-342.)
- [9] Zelniker E E, Hospedales T M, Gong S G, et al. A unified bayesian framework for adaptive visual tracking[C]. *British Machine Vision Conf*. London: British Machine Vision Association, 2009: 1-9.
- [10] Kwon J, Lee K M. Visual tracking decomposition[C]. *Proc of IEEE Conf on Computer Vision and Pattern Recognition*. San Francisco: IEEE Press, 2010: 1269-1276.
- [11] Qian Y, Thang B D, Gerard M. Online tracking and reacquisition using co-trained generative and discriminative Trackers[C]. *Proc of the 10th European Conf on Computer Vision: Part II*. Marseille: Springer-Verlag, 2008: 678-691.
- [12] Avidan S. Support vector tracking[J]. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 2004, 26(8): 1064-1072.
- [13] Doucet A, Godsill S, Andrieu C. On sequential monte carlo sampling methods for bayesian filtering[J]. *Statistics and Computing*, 2000, 10(3): 197-208.
- [14] Liu B, Huang J, Kulikowski C, et al. Robust visual tracking using local sparse appearance model and k -selection[J]. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 2013, 35(12): 2968-2981.
- [15] Babenko B, Yang M H, Belongie S. Visual tracking with online multiple instance learning[C]. *Proc of IEEE Conf on Computer Vision and Pattern Recognition*. Miami: IEEE Press, 2009: 983-990.
- [16] Zhang K H, Zhang L, Yang M H, et al. Fast tracking via spatio-temporal context learning[C]. *Proc of the 13th European Conf on Computer Vision: Part V*. Zurich: Springer International Publishing, 2014: 127-141.
- [17] Zhuang B, Lu H, Xiao Z, et al. Visual tracking via discriminative sparse similarity map[J]. *IEEE Trans on Image Processing*, 2014, 23(4): 1872-1881.

(责任编辑: 闫 妍)