

基于代价敏感结构化SVM的目标跟踪

袁广林* 孙子文 秦晓燕 夏良 朱虹

(中国人民解放军陆军炮兵防空兵学院信息工程系 合肥 230031)

摘要: 基于结构化SVM的目标跟踪由于其优异的性能而受到了广泛关注,但是现有方法存在正样本和负样本不平衡问题。针对此问题,该文首先提出一种用于目标跟踪的代价敏感结构化SVM模型,其次基于对偶坐标下降原理设计了该模型的求解算法,最后利用提出的代价敏感结构化SVM实现了一种多尺度目标跟踪方法。在OTB100数据集和VOT2019数据集上进行了实验验证,实验结果表明:该文方法相比相关滤波目标跟踪方法,跟踪精度较高,相比深度目标跟踪方法,具有速度优势。

关键词: 目标跟踪; 非平衡问题; 代价敏感; 结构化SVM

中图分类号: TN911.73; TP391.4

文献标识码: A

文章编号: 1009-5896(2021)11-3335-07

DOI: 10.11999/JEIT200708

Object Tracking Based on Cost Sensitive Structured SVM

YUAN Guanglin SUN Ziwen QIN Xiaoyan XIA Liang ZHU Hong

(Department of Information Engineering of PLA Army Academy of Artillery
and Air Defense, Hefei 230031, China)

Abstract: Object tracking based on structured SVM attracts much attention due to its excellent performance. However, the existing methods have the problem of imbalance between positive and negative samples. To solve the problem, a cost sensitive structured SVM model is proposed for object tracking. Secondly, an algorithm for the proposed model is designed via dual coordinate descent principle. Finally, a multi-scale object tracking method is implemented using the proposed cost sensitive structured SVM. The experimental results on OTB100 datasets and VOT2019 datasets show that compared with the correlation filtering trackers, the proposed method has higher tracking accuracy, and has the advantage of speed compared with the deep object trackers.

Key words: Object tracking; Unbalanced problem; Cost sensitive; Structured SVM

1 引言

目标跟踪是计算机视觉领域中的基础研究课题,它是视频分析中的一项重要技术,其目标是利用视频数据估计目标的状态。目标跟踪在视频监控、车辆导航、人机交互、智能交通、运动分析和姿态估计等民用领域,以及视觉制导、目标定位和火力控制等军事领域均有重要的应用价值。近年来,虽然目标跟踪有了较大发展,但是其仍然面临复杂背景、目标变化和快速运动等诸多难题,目前仍然是计算机视觉领域中研究的热点问题。

受到结构化SVM在目标检测中应用的启发,Hare等人^[1]于2011年在ICCV上首次提出基于结构化SVM的目标跟踪方法Struck,2015年该文的扩

展^[2]发表在顶级国际期刊*IEEE Transactions on Pattern Analysis and Machine Intelligence*上。Struck把目标跟踪看作结构化学习问题,避免了传统判别式跟踪的中间分类环节,显著提高了目标跟踪的性能。为了适应目标的变化同时又不丢失目标的时间上下文信息,Yao等人^[3]于2012年提出一种加权在线结构化SVM跟踪方法,进一步提升了结构化SVM跟踪的性能。为了提高遮挡与形变目标的跟踪性能,2013年,Yao等人^[4]又以在线算法Pegasos为基础,提出一种在线结构化SVM用于目标跟踪。为了解决目标跟踪中存在的模型漂移问题,2012年Bai等人^[5]提出一种在线拉普拉斯排序SVM跟踪。同样是为了应对模型漂移问题,Zhang等人^[6]于2014年提出了MEEM跟踪,该方法以在线SVM作为基础跟踪建立专家组跟踪目标,取得了较好的效果。2015年Hong等人^[7]利用在线SVM指导反向传播特定目标的CNN特征到输入层,进而建立特定目标的显著图跟踪目标,该方法建立的显

收稿日期:2020-08-10; 改回日期:2021-04-14; 网络出版:2021-07-11

*通信作者:袁广林 yuangl_plus@126.com

基金项目:安徽省自然科学基金(2008085QF325)

Foundation Item: Anhui Provincial Natural Science Foundation (2008085QF325)

著图保持了目标的空间结构,增强了目标跟踪的鲁棒性。2016年Ning等人^[8]基于对偶坐标下降原理提出一种对偶线性结构化SVM目标跟踪方法,该方法保证了目标跟踪的鲁棒性,同时又提高了目标跟踪的速度。2017年Wang等人^[9]提出LMCF跟踪,该方法利用相关滤波对基于结构化SVM的目标跟踪进行了加速。2018年Ji等人^[10]采用了与LMCF跟踪类似的思想来加速基于目标部件的结构化SVM跟踪。2019年Zuo等人^[11]提出一种基于离散傅里叶变换的交替优化方法求解相关滤波器,并设计了多通道支撑相关滤波器跟踪目标,进一步提升了基于SVM的目标跟踪的性能。

综上所述,基于结构化SVM的目标跟踪方法具有较优的跟踪性能,受到了广泛的关注,但是现有方法存在正样本和负样本不平衡的问题。针对基于结构化SVM的目标跟踪中存在的负样本和正样本不平衡的问题,本文提出一种代价敏感结构化SVM模型,基于对偶坐标优化原理设计了该模型的求解算法,并利用该算法实现了单尺度目标跟踪算法(Dual Liner Cost Sensitive Structured Support Machine, DLCS-SSVM)和多尺度目标跟踪方法(Scale Dual Liner Cost Sensitive Structured Support Machine, Scale-DLCS-SSVM)。利用OTB100数据集^[12]和VOT2019数据集^[13]对提出的目标跟踪方法进行了实验验证,并与现有的先进目标跟踪方法进行了比较。实验结果表明,本文提出的目标跟踪方法达到了预期的跟踪效果,与现有目标跟踪方法相比具有较好的性能。

2 代价敏感和结构化支持向量机

2.1 代价敏感支持向量机

为了解决SVM中存在的正负样本不平衡的问题,文献^[14]提出一种代价敏感支持向量机。假设二分类训练数据集是 $\{(\mathbf{x}_i, y_i)\}_{i=1}^N$,其中 $\mathbf{x}_i \in \mathbf{R}^d$ 是样本特征向量, $y_i \in +1, -1$ 是类别标签,则文献^[14]提出的代价敏感支持向量机描述为

$$\begin{aligned} & \operatorname{argmin}_{\mathbf{w}, b, \xi_i} \frac{1}{2} \|\mathbf{w}\|^2 + C \left(C_1 \sum_{\{i|y_i=1\}} \xi_i + \frac{1}{\kappa} \sum_{\{i|y_i=-1\}} \xi_i \right) \\ & \forall \xi_i \geq 0 \\ & \text{s.t. } (\mathbf{w}^T \mathbf{x}_i + b) \geq 1 - \xi_i; y_i = +1 \\ & (\mathbf{w}^T \mathbf{x}_i + b) \leq -\kappa + \xi_i; y_i = -1 \end{aligned} \quad (1)$$

其中

$$\kappa = \frac{1}{2C_1 - 1}, 0 < \kappa \leq 1 \leq \frac{1}{\kappa} \leq C_1 \quad (2)$$

\mathbf{w} 和 b 分别是SVM分类器的法向量和偏置, ξ_i 是松弛变量, C, C_1, C_1 和 κ 是正则化参数。

2.2 结构化支持向量机

受到结构化SVM在目标检测中应用的启发,2011年Hare等人^[1]提出了基于结构化SVM的目标跟踪Struck。Struck把目标跟踪看作结构化学习问题,避免了传统判别式跟踪的中间分类环节,显著提高了目标跟踪的性能。在利用结构化学习器进行预测时,其目标是预测给定样本 $\mathbf{x} \in \mathbf{R}^d$ 的结构化输出 $\mathbf{y} \in \mathbf{Y}$,其中 \mathbf{Y} 可以是任意结构输出空间。在基于结构化SVM的目标跟踪中, \mathbf{Y} 是矩形框空间,它的任一元素用 (x, y, w, h) 表示,其中 (x, y) 表示矩形框的中心位置, w 和 h 分别表示矩形框的宽和高。假设训练数据为 $\{(\mathbf{x}_i, \mathbf{y}_i)\}_{i=1}^N$, $\Phi(\mathbf{x}, \mathbf{y})$ 表示矩形框 \mathbf{y} 在图像 \mathbf{x} 上的特征向量,则可以通过式(3)学习一个参数为 \mathbf{w} 的分类器。

$$\mathbf{y}^* = \operatorname{argmax}_{\mathbf{y} \in \mathbf{Y}} f(\mathbf{x}, \mathbf{y}, \mathbf{w}) \quad (3)$$

其中, $f(\mathbf{x}, \mathbf{y}, \mathbf{w}) = \mathbf{w}^T \Phi(\mathbf{x}, \mathbf{y})$,分类器法向量 \mathbf{w} 通过求解如下优化问题得到。

$$\begin{aligned} & \min_{\mathbf{w}} \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^N \xi_i \\ & \text{s.t. } \forall i, \xi_i \geq 0 \\ & \forall i, \forall \mathbf{y} \neq \mathbf{y}_i : \langle \mathbf{w}, \Psi_i(\mathbf{y}) \rangle \geq L(\mathbf{y}_i, \mathbf{y}) - \xi_i \end{aligned} \quad (4)$$

其中, $\Psi_i(\mathbf{y}) = \Phi(\mathbf{x}_i, \mathbf{y}_i) - \Phi(\mathbf{x}_i, \mathbf{y})$, C 是正则化参数, ξ_i 是松弛变量。 $L(\mathbf{y}_i, \mathbf{y})$ 是损失函数,表示预测输出矩形框 \mathbf{y} 的结构误差损失。 $L(\mathbf{y}_i, \mathbf{y})$ 由样本矩形 \mathbf{y}_i 和目标矩形 \mathbf{y} 的重合率定义,如式(5)所示。

$$L(\mathbf{y}_i, \mathbf{y}) = 1 - s(\mathbf{y}_i, \mathbf{y}) \quad (5)$$

其中

$$s(\mathbf{y}_i, \mathbf{y}) = \frac{\mathbf{y}_i \cap \mathbf{y}}{\mathbf{y}_i \cup \mathbf{y}} \quad (6)$$

3 代价敏感结构化SVM

3.1 代价敏感结构化SVM模型

现有基于结构化SVM的目标跟踪方法存在正样本和负样本不平衡的问题。如图1所示,与目标重叠区域较大的训练样本称为正样本,与目标重叠区域较小的训练样本称为负样本。目标跟踪需要在当前目标周围区域中进行采样更新表观模型。从图1可以看出,在采样得到的样本中,负样本的数量远大于正样本的数量,即负样本与正样本严重不平衡。其中,黑色实线是非代价敏感结构化SVM的超平面,红色虚线是代价敏感结构化SVM的超平面。当训练数据集不平衡时,SVM分类器对少数类的识别率较低,因此该问题制约了目标跟踪的性能。为了解决正样本和负样本不平衡对基于结构化SVM目标跟踪方法性能的影响,本文将文献^[14]的思想引入到文

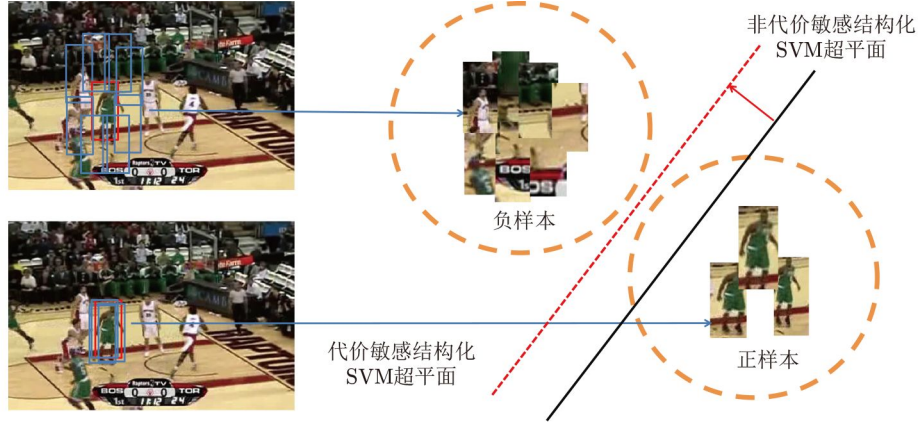


图1 结构化SVM目标跟踪中存在的正样本和负样本不平衡问题

献[1]提出的结构化SVM中, 设计一种基于代价敏感的结构化SVM模型, 描述为

$$\min_{\mathbf{w}} \frac{1}{2} \|\mathbf{w}\|^2 + C \left(C^+ \sum_{\substack{\mathbf{y} \neq \mathbf{y}_i, \\ L(\mathbf{y}_i, \mathbf{y}) < m}} \xi_i + \frac{1}{\kappa} \sum_{\substack{\mathbf{y} \neq \mathbf{y}_i, \\ L(\mathbf{y}_i, \mathbf{y}) \geq m}} \xi_i \right)$$

$$\forall i, \xi_i \geq 0$$

$$\text{s.t. } \forall i, \forall \mathbf{y} \neq \mathbf{y}_i, L(\mathbf{y}_i, \mathbf{y}) < m$$

$$\langle \mathbf{w}, \Psi_i(\mathbf{y}) \rangle \geq L(\mathbf{y}_i, \mathbf{y}) - \xi_i$$

$$\forall i, \forall \mathbf{y} \neq \mathbf{y}_i, L(\mathbf{y}_i, \mathbf{y}) \geq m$$

$$\langle \mathbf{w}, \Psi_i(\mathbf{y}) \rangle \geq \kappa L(\mathbf{y}_i, \mathbf{y}) - \xi_i \quad (7)$$

其中,

$$\kappa = \frac{1}{2C^+ - 1}, \quad 0 < \kappa \leq 1 \leq \frac{1}{\kappa} \leq C^+ \quad (8)$$

$\Psi_i(\mathbf{y}) = \Phi(\mathbf{x}_i, \mathbf{y}_i) - \Phi(\mathbf{x}_i, \mathbf{y})$, $L(\mathbf{y}_i, \mathbf{y}) = 1 - s(\mathbf{y}_i, \mathbf{y})$, $s(\mathbf{y}_i, \mathbf{y}) = \frac{\mathbf{y}_i \cap \mathbf{y}}{\mathbf{y}_i \cup \mathbf{y}}$. C^+ , C 和 κ 是正则化参数. m 是阈值. 当式(7)中损失函数 $L(\mathbf{y}_i, \mathbf{y})$ 的值大于等于阈值 m 时, 训练样本 $(\mathbf{x}_i, \mathbf{y}_i)$ 定义为负样本; 否则当 $L(\mathbf{y}_i, \mathbf{y})$ 值小于阈值 m 时, 训练样本 $(\mathbf{x}_i, \mathbf{y}_i)$ 定义为正样本.

3.2 代价敏感结构化SVM算法

为了求解本文提出的代价敏感结构化支持向量机式(7), 使用拉格朗日乘子法可得到其对偶问题, 为此引入拉格朗日乘子 $\alpha_i^{\mathbf{y}}$ 和 $\beta_i^{\mathbf{y}}$, 满足条件

$$\forall i, \forall \mathbf{y} \neq \mathbf{y}_i : \alpha_i^{\mathbf{y}} \geq 0, \beta_i^{\mathbf{y}} \geq 0 \quad (9)$$

则式(7)的拉格朗日函数为

$$L(\mathbf{w}, \xi, \alpha, \beta) = \frac{1}{2} \|\mathbf{w}\|^2$$

$$+ C \left(C^+ \sum_{\substack{i, \mathbf{y} \neq \mathbf{y}_i \\ L(\mathbf{y}_i, \mathbf{y}) < m}} \xi_i + \frac{1}{\kappa} \sum_{\substack{i, \mathbf{y} \neq \mathbf{y}_i \\ L(\mathbf{y}_i, \mathbf{y}) \geq m}} \xi_i \right)$$

$$- \sum_{\substack{i, \mathbf{y} \neq \mathbf{y}_i \\ L(\mathbf{y}_i, \mathbf{y}) < m}} \alpha_i^{\mathbf{y}} (\langle \mathbf{w}, \Psi_i(\mathbf{y}) \rangle + \xi_i - L(\mathbf{y}_i, \mathbf{y}))$$

$$- \sum_{\substack{i, \mathbf{y} \neq \mathbf{y}_i \\ L(\mathbf{y}_i, \mathbf{y}) \geq m}} \alpha_i^{\mathbf{y}} (\langle \mathbf{w}, \Psi_i(\mathbf{y}) \rangle + \xi_i - \kappa L(\mathbf{y}_i, \mathbf{y}))$$

$$- \sum_{i, \mathbf{y} \neq \mathbf{y}_i} \beta_i^{\mathbf{y}} \xi_i \quad (10)$$

将拉格朗日函数 $L(\mathbf{w}, \xi, \alpha, \beta)$ 分别对 \mathbf{w} 和 ξ_i 求偏导并令其为0可得

$$\mathbf{w} = \sum_{i, \mathbf{y} \neq \mathbf{y}_i} \alpha_i^{\mathbf{y}} \Psi_i(\mathbf{y}) \quad (11)$$

$$\forall i, \mathbf{y} \neq \mathbf{y}_i, L(\mathbf{y}_i, \mathbf{y}) < m, CC^+ - \alpha_i^{\mathbf{y}} - \beta_i^{\mathbf{y}} = 0 \quad (12)$$

$$\forall i, \mathbf{y} \neq \mathbf{y}_i, L(\mathbf{y}_i, \mathbf{y}) \geq m, \frac{C}{\kappa} - \alpha_i^{\mathbf{y}} - \beta_i^{\mathbf{y}} = 0 \quad (13)$$

将式(11)–式(13)代入式(10), 可以将 $L(\mathbf{w}, \xi, \alpha, \beta)$ 中的 \mathbf{w} , β 和 ξ 消去, 得到式(7)的对偶问题, 如式14(a)–式14(c)所示.

$$\min_{\alpha \geq 0} L(\alpha) = \frac{1}{2} \left\| \sum_{i, \mathbf{y} \neq \mathbf{y}_i} \alpha_i^{\mathbf{y}} \Psi_i(\mathbf{y}) \right\|^2 - \sum_{\substack{i, \mathbf{y} \neq \mathbf{y}_i, \\ L(\mathbf{y}_i, \mathbf{y}) < m}} \alpha_i^{\mathbf{y}} L(\mathbf{y}_i, \mathbf{y})$$

$$- \kappa \sum_{\substack{i, \mathbf{y} \neq \mathbf{y}_i, \\ L(\mathbf{y}_i, \mathbf{y}) \geq m}} \alpha_i^{\mathbf{y}} L(\mathbf{y}_i, \mathbf{y}) \quad (14a)$$

$$\text{s.t. } \forall i, \mathbf{y} \neq \mathbf{y}_i, L(\mathbf{y}_i, \mathbf{y}) < m, \alpha_i^{\mathbf{y}} \geq 0, \sum_{\mathbf{y}} \alpha_i^{\mathbf{y}} \leq CC^+ \quad (14b)$$

$$\forall i, \mathbf{y} \neq \mathbf{y}_i, L(\mathbf{y}_i, \mathbf{y}) \geq m, \alpha_i^{\mathbf{y}} \geq 0, \sum_{\mathbf{y}} \alpha_i^{\mathbf{y}} \leq \frac{C}{\kappa} \quad (14c)$$

本文模型采用线性核, 并基于对偶坐标优化原理^[15]设计模型式(14)的求解算法. 由文献[15]可知, 对偶坐标优化(Dual Coordinate Descent, DCD)算法每次利用式(15)从训练集中选择一个训练样本 k , 然后利用式(16)更新其对偶标量 $\alpha_k^{\mathbf{y}}$.

$$\mathbf{y}_k^* = \arg \max_{\mathbf{y} \in \mathbf{Y}_k} L(\mathbf{y}, \mathbf{y}_k) - \mathbf{w}^T \Psi_k(\mathbf{y}) \quad (15)$$

$$\alpha_k^{\mathbf{y}^* \text{(new)}} = \alpha_k^{\mathbf{y}^* \text{(old)}} + \Delta \alpha_k^{\mathbf{y}^*} \quad (16)$$

解决问题的关键是如何求得式(16)中的 $\Delta \alpha_k^{\mathbf{y}^*}$. 为此, 首先将式(16)代入式(14a), 然后将其转化为关于 $\Delta \alpha_k^{\mathbf{y}^*}$ 的函数, 得到式(17).

$$\begin{aligned}
L(\Delta\alpha_k^{y^*}) = & \frac{1}{2} \left(\alpha_k^{y^* (\text{old})} + \Delta\alpha_k^{y^*} \right)^2 \|\Psi_k(\mathbf{y}^*)\|^2 \\
& + \Delta\alpha_k^{y^*} \sum_{\substack{i \neq k \\ \mathbf{y} \neq \mathbf{y}_i^* \\ \mathbf{y} \neq \mathbf{y}^*}} \alpha_i^{y^*} \Psi_i(\mathbf{y}) \Psi_k(\mathbf{y}) \\
& - \Delta\alpha_k^{y^*} L(\mathbf{y}_k, \mathbf{y}^*) \Big|_{L(\mathbf{y}_k, \mathbf{y}_i) < m} \\
& - \kappa \Delta\alpha_k^{y^*} L(\mathbf{y}_k, \mathbf{y}^*) \Big|_{L(\mathbf{y}_k, \mathbf{y}_i) \geq m} + c \quad (17)
\end{aligned}$$

其中, c 是一个与增量 $\Delta\alpha_k^{y^*}$ 无关的常量。将式(17)对 $\Delta\alpha_k^{y^*}$ 求导并令其为0可得

$$\begin{aligned}
& \forall i, \mathbf{y} \neq \mathbf{y}_i, L(\mathbf{y}_i, \mathbf{y}) < m, \\
\Delta\alpha_k^{y^*} = & \frac{L(\mathbf{y}_k, \mathbf{y}^*) - \mathbf{w}^T \Psi_k(\mathbf{y}^*)}{\|\Psi_k(\mathbf{y}^*)\|^2} \quad (18)
\end{aligned}$$

$$\begin{aligned}
& \forall i, \mathbf{y} \neq \mathbf{y}_i, L(\mathbf{y}_i, \mathbf{y}) \geq m, \\
\Delta\alpha_k^{y^*} = & \frac{\kappa L(\mathbf{y}_k, \mathbf{y}^*) - \mathbf{w}^T \Psi_k(\mathbf{y}^*)}{\|\Psi_k(\mathbf{y}^*)\|^2} \quad (19)
\end{aligned}$$

根据约束条件式(14b)可得

当 $L(\mathbf{y}_i, \mathbf{y}) < m$ 时, $\Delta\alpha_k^{y^*}$ 的取值范围为

$$\left[-\alpha_k^{y^* (\text{old})}, CC^+ - \sum_{\mathbf{y}} \alpha_k^{\mathbf{y}} \right] \quad (20)$$

根据约束条件式14(c)可得

当 $L(\mathbf{y}_i, \mathbf{y}) \geq m$ 时, $\Delta\alpha_k^{y^*}$ 的取值范围为

$$\left[-\alpha_k^{y^* (\text{old})}, \frac{C}{\kappa} - \sum_{\mathbf{y}} \alpha_k^{\mathbf{y}} \right] \quad (21)$$

利用式(18)—式(21)可以得到 $\Delta\alpha_k^{y^*}$, 然后结合式(11)即可得到 \mathbf{w} 的更新公式, 如式(22)所示。

$$\mathbf{w}^{(\text{new})} = \mathbf{w}^{(\text{old})} + \Delta\alpha_k^{y^*} \Psi_k(\mathbf{y}^*) \quad (22)$$

在基于结构化SVM的目标跟踪中, 随着时间的推移, 结构化SVM中的支持向量的数量不断增加。为了保证目标跟踪的效率, 需要固定支持向量的数目。为此, 当结构化SVM中模式数超出预算时, 根据式(23)选择一个支持向量删除, 本文提出的代价敏感结构化SVM采用这一策略。

$$\alpha^* = \arg \min_{\alpha_i^{\mathbf{y}} \in \alpha} \|\alpha_i^{\mathbf{y}} \Psi_i(\mathbf{x}_i, \mathbf{y})\|^2 \quad (23)$$

综上所述, 本文提出的代价敏感结构化SVM学习算法如算法1所示。

算法1 代价敏感结构化SVM学习算法

输入: $t-1$ 时刻代价敏感结构化SVM的参数 \mathbf{w}_{t-1} 和模式集 \mathbf{P}_{t-1} , t 时刻的训练样本 $\{\mathbf{x}_i^i, \mathbf{y}_i^i\}_{i=1}^N$ 。

输出: t 时刻代价敏感结构化SVM的参数 \mathbf{w}_t 和模式集 \mathbf{P}_t 。

步骤1 $\mathbf{P}_t = \mathbf{P}_{t-1} \cup \{\mathbf{x}_i^i, \mathbf{y}_i^i\}_{i=1}^N$;

For $j=1: n_1$

步骤2 计算模式集 \mathbf{P}_t 中的模式数 n ;

步骤3 令 $i = n - \lfloor (j-1) * n/n_1 \rfloor$;

步骤4 用式(15), 从模式 \mathbf{P}_t^i 中选择样本;

步骤5 用式(16), 更新 \mathbf{y}_i^* 的对偶变量 $\alpha_i^{y^*}$;

步骤6 用式(22), 更新 \mathbf{w}_{t-1} ;

步骤7 如果模式集 \mathbf{P}_t 中的支撑向量数目大于预算数目, 则用式(23)更新模式集 \mathbf{P}_t ;

步骤8 计算模式集 \mathbf{P}_t 中的模式数 n ;

For $p=1: n_2$

令 $i = n - \lfloor (p-1) * n/n_2 \rfloor$;

步骤9 用式(15), 从模式 \mathbf{P}_t^i 中选择样本 \mathbf{y}_i^* ;

步骤10 用式(16), 更新 \mathbf{y}_i^* 的对偶变量 $\alpha_i^{y^*}$;

步骤11 用式(22), 更新 \mathbf{w}_{t-1} ;

End For

End For

步骤12 令 $\mathbf{w}_t = \mathbf{w}_{t-1}$ 。

说明: 算法1中 n_1 和 n_2 是外部循环和内循环的迭代次数, 本文分别取5和10。

4 目标跟踪方法

本文使用网格搜索生成候选样本, 搜索区域的大小由跟踪目标的大小自适应确定。样本的大小设定为 20×20 。在得到代价敏感结构化SVM的超平面 \mathbf{w} 后, 利用内积运算计算候选样本的得分, 根据最大得分准则式(24)估计目标的状态。

$$\mathbf{y}^* = \arg \max_{\mathbf{y} \in \mathbf{Y}} \mathbf{w}^T \Psi_t(\mathbf{y}) \quad (24)$$

对于目标特征, 本文选择目标的Lab颜色和局部秩变换(Local Rank Transformation, LRT)特征, LRT特征的计算方法同文献[16], 这里不再详述。基于上面的分析, 本文提出的代价敏感结构化SVM目标跟踪方法描述如下。

方法1 代价敏感结构化SVM目标跟踪方法

输入: 序列图像 $I = \{I_1, I_2, \dots, I_T\}$ 和目标状态 \mathbf{s}_1 , 尺度参数 $\text{Sp} = \{\text{Sp}[1], \text{Sp}[2], \dots, \text{Sp}[n]\}$

输出: 每一帧的跟踪结果 \mathbf{s}_i ($i = 1, 2, \dots, T$)。

步骤1 根据当前目标状态 \mathbf{s}_1 , 采样 N 个训练样本 $\{(\mathbf{x}_1^i, \mathbf{y}_1^i)\}_{i=1}^N$;

步骤2 初始 $\mathbf{w}_0 = \mathbf{0}$ 和 $\mathbf{P}_0 = \mathbf{0}$;

步骤3 调用算法1初始化 $t=1$ 时刻 DLCS_SVM 的参数 \mathbf{w}_1 和模式 \mathbf{P}_1 ;

For ($t=2: T$),

For ($k=1: \text{length}(\text{Sp})$)

$I_{t,k} = \text{Scale}(I_t, \text{Sp}[k])$;

步骤4 利用滑动窗口在 $I_{t,k}$ 上的搜索区域中采样测试样本 $\{(\mathbf{x}_{t,k}^i, \mathbf{y}_{t,k}^i)\}_{i=1}^M$, $\mathbf{X}_{t,k} = [\mathbf{x}_{t,k}^i]_{i=1}^M$, $\mathbf{Y}_{t,k} = [\mathbf{y}_{t,k}^i]_{i=1}^M$;

步骤5 利用最大化得分准则估计目标当前状

态: $\mathbf{s}_{t,k}^m = \arg \max_{\mathbf{y} \in \mathbf{Y}_{t,k}} (\mathbf{w}_t^T \mathbf{X}_{t,k});$

End For

$\mathbf{s}_t = \max(\mathbf{s}_{t,k}^m), m \in Sp$

步骤6 根据当前目标状态 \mathbf{s}_t , 采样 N 个训练样本 $\{(\mathbf{x}_t^i, \mathbf{y}_t^i)\}_{i=1}^N$;

步骤7 调用算法1更新DLCS_SSVM的 \mathbf{w}_t ;

End For

方法1中 $I_{t,k} = \text{Scale}(I_t, Sp[k])$ 的功能是利用尺度参数 $Sp[k]$ 对图像 I_t 进行缩放, 结果赋予 $I_{t,k}$ 。对于单尺度目标跟踪, 尺度参数 Sp 设定为 $\{1\}$, 即可完成单尺度目标跟踪, 称为单尺度代价敏感结构化SVM目标跟踪方法(简称DLCS_SSVM)。对于多尺度目标跟踪, 尺度参数 $Sp = \{Sp[1], Sp[2], \dots, Sp[n]\}$, 本文 $Sp = \{1, 0.995, 1.005\}$, 即在3种不同尺度图像上分别跟踪目标, 以最大得分作为跟踪结果, 即可完成多尺度目标跟踪, 称为多尺度代价敏感结构化SVM目标跟踪(Scale-DLCS_SSVM)。

5 实验

5.1 实验环境、数据集和参数设置

在SYS-7048GR-TR台式机(CPU型号为Intel Xeon(R) ES-2630v4@2.20 GHz×20, 内存为64 GB, GPU为RTX2080Ti 11 GB)上使用Matlab和OpenCV实现了本文提出的跟踪方法, 其中Matlab版本为R2017a, OpenCV版本为2.4.8。一方面利用OTB100数据集^[12]对提出的目标跟踪方法进行实验验证。评价指标为OPE, TRE和SRE^[12]。另一方面利用VOT2019数据集^[13]对提出的目标跟踪方法进行实验验证。评价指标为EAO, Accuracy, Robustness^[13]。本文提出的目标跟踪方法中有一些需要设置参数, 实验中这些参数固定不变。惩罚系数 C 的值为100, 支持向量的预算设为100, C^+ 的值设为2, κ 的值设为0.67, 正样本与负样本的阈值 m 设为0.5。多尺度估计参数设置为 $Scale = \{1, 0.995, 1.005\}$ 。

5.2 实验结果及分析

5.2.1 与其它基于结构化SVM跟踪器在OTB100数据集上的比较

表1给出了本文方法DLCS_SSVM, Scale-DLCS_SSVM, DLSSVM^[8], Scale-DLSSVM^[8], Struck^[1]和LMCF^[9]等6种基于结构化SVM跟踪器在OTB100数据集上的OPE性能指标比较结果。从表1可以看出, 在精度和成功率两项指标上, 本文提出的方法Scale-DLCS_SSVM均明显优于其它跟踪器。在跟踪速度上, 本文选择一个长视频liquor(1741帧)^[12]进行评估。由表1中的比较结果可以看出: 本文方法与DLSSVM^[8]方法相比不仅在性能上有明显提升, 而且对目标跟踪速度几乎没有影响。

5.2.2 与其它高性能跟踪器跟踪速度比较

由表2与其他高性能跟踪器跟踪速度比较可知, 相较于深度学习目标跟踪方法, 如DeepLMCF^[9]和DeepSRDCF^[17], 本文方法速度明显更快。相较于相关滤波目标跟踪方法, 如TADT^[18], 本文单尺度方法速度与其速度相当, 能够达到实时的跟踪效果。

5.2.3 与其它高性能跟踪器在OTB100数据集上的比较

在OTB100数据集上, 将本文提出的目标跟踪方法与4种优秀的目标跟踪方法进行比较。4种方法分别是基于深度学习的DeepSRDCF^[17]、基于相关滤波的Staple^[19]、基于结构化SVM的多尺度跟踪Scale-DLSSVM^[8]及基于结构化SVM与相关滤波的LMCF^[9]。图2为6种跟踪器在OTB100数据集上取效果前5名的OPE, TRE和SRE性能指标曲线。从图2中的结果可以看出, 本文提出的Scale-DLCS_SSVM与DLCS_SSVM跟踪器在准确度和成功率两个指标上相比Scale-DLSSVM跟踪方法都有明显的提高。且在OPE评价指标上Scale-DLCS_SSVM相比DeepSRDCF^[17]在成功率上高1.3%, 在TRE评价指标上Scale-DLCS_SSVM相比DeepSRDCF^[17]在成功率上高1.5%, 在SRE评价指标上Scale-

表1 6种基于结构化SVM的跟踪器在OTB100数据集上的OPE性能与速度指标

跟踪方法	精度(pixels)	成功率(AUC)	速度(fps)
Scale-DLCS_SSVM	0.837	0.648	12.18
DLCS_SSVM	0.805	0.603	25.32
Scale-DLSSVM ^[8]	0.803	0.561	11.47
DLSSVM ^[8]	0.765	0.540	23.16
Struck ^[1]	0.635	0.459	7.52
LMCF ^[9]	0.786	0.578	78.23

DLCS_SVM相比DeepSRDCF^[17]在成功率上高1.0%。

5.2.4 与其它高性能跟踪器在VOT2019数据集上的比较

如表3所示，本文选取近两年高性能的深度学习与相关滤波跟踪方法在VOT2019数据集上进行比较。由表3可知，本文多尺度方法在单个指标上略低于SiamMask，但是相较于其他高性能方法来说在各个性能上有着优势。

从上述所有实验结果可以看出：本文提出的目标跟踪方法达到了预期的跟踪效果，其跟踪性能优于现有基于结构化SVM的目标跟踪方法；与相关滤波目标跟踪方法相比，本文方法跟踪精度较高；与深度学习目标跟踪方法相比，本文方法具有速度优势。

6 结束语

本文分析了基于结构化SVM的目标跟踪在进行训练时存在正样本和负样本不平衡的问题。针对该问题，基于代价敏感SVM和结构化SVM提出一种代价敏感结构化SVM模型，并利用对偶坐标下降优化设计了代价敏感结构化SVM算法。最后，利用提出的代价敏感结构化SVM实现了一种多尺度目标跟踪方法。利用OTB100数据集和VOT2019数据集分别对提出的目标跟踪方法进行了实验验证和分析。实验结果表明，与相关滤波目标跟踪中的一些优秀方法相比，本文方法跟踪精度较高，与深度目标跟踪中的一些优秀方法相比，本文方法具有速度优势。在目标跟踪中，由于跟踪误差会引起模型漂移，进而导致跟踪失败，如何利用结构化SVM解决这一问题进一步研究的方向。

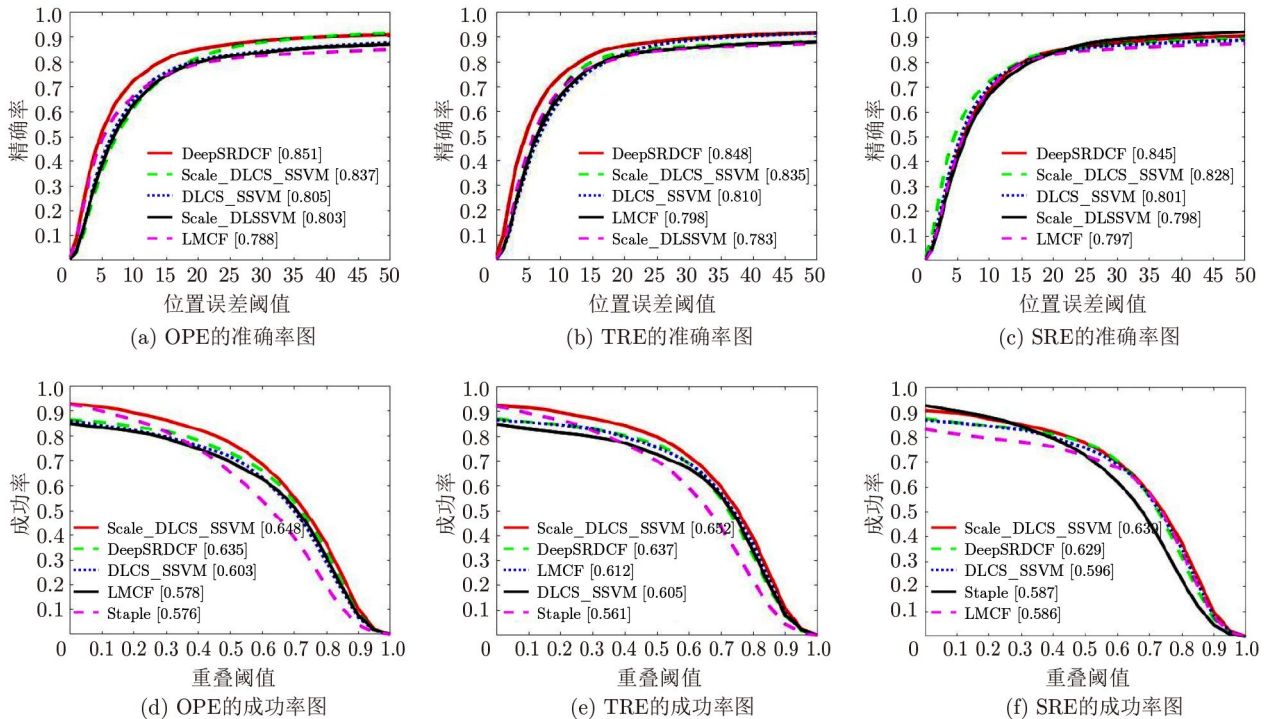


图 2 6种高性能跟踪器在OTB100数据集上取效果前5名的OPE, TRE和SRE性能指标曲线

表 2 5种跟踪方法在OTB100数据集上的速度指标(fps)

跟踪方法	Scale-DLCS_SVM	DLCS_SVM	DeepLMCF ^[9]	DeepSRDCF ^[17]	TADT ^[18]
速度	12.18	25.32	7.45	1.56	27.89

表 3 本文多尺度方法与近年来4种高性能跟踪方法在VOT2019数据集上实验结果

	本文	SiamMask ^[20]	GradNet ^[21]	TADT ^[18]	DeepSRDCF ^[17]
EAO ↑	0.217	0.289	0.213	0.207	0.198
Accuracy ↑	0.538	0.596	0.531	0.516	0.512
Robustness ↓	0.559	0.461	0.538	0.567	0.572

参考文献

- [1] HARE S, SAFFARI A, and TORR P H S. Structured output tracking with kernels[C]. 2011 International Conference on Computer Vision, Barcelona, Spain, 2011: 263-270.
- [2] HARE S, GOLODETZ S, SAFFARI A, *et al.* Structured output tracking with kernels[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, 38(10): 2096-2109. doi: [10.1109/TPAMI.2015.2509974](https://doi.org/10.1109/TPAMI.2015.2509974).
- [3] YAO Rui, SHI Qinfeng, SHEN Chunhua, *et al.* Robust tracking with weighted online structured learning[C]. The 12th European Conference on Computer Vision, Florence, Italy, 2012: 158-172.
- [4] YAO Rui, SHI Qinfeng, SHEN Chunhua, *et al.* Part-based visual tracking with online latent structural learning[C]. 2013 IEEE Conference on Computer Vision and Pattern Recognition, Portland, USA, 2013: 2363-2370.
- [5] BAI Yancheng and TANG Ming. Robust tracking via weakly supervised ranking SVM[C]. 2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, USA, 2012: 1854-1861.
- [6] ZHANG Jianming, MA Shugao, and SCLAROFF S. MEEM: Robust tracking via multiple experts using entropy minimization[C]. The 13th European Conference on Computer Vision, Zurich, Swiss, 2014: 188-203.
- [7] HONG S, YOU T, KWAK S, *et al.* Online tracking by learning discriminative saliency map with convolutional neural network[C]. The 32nd International Conference on International Conference on Machine Learnin, Lille, France, 2015: 597-606.
- [8] NING Jifeng, YANG Jimei, JIANG Shaojie, *et al.* Object tracking via dual linear structured SVM and explicit feature map[C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, USA, 2016: 4266-4274.
- [9] WANG Mengmeng, LIU Yong, and HUANG Zeyi. Large margin object tracking with circulant feature maps[C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, USA, 2017: 4800-4808.
- [10] JI Zhangjian, FENG Kai, and QIAN Yuhua. Part-based visual tracking via structural support correlation filter[J]. *Journal of Visual Communication and Image Representation*, 2019, 64: 102602. doi: [10.1016/j.jvcir.2019.102602](https://doi.org/10.1016/j.jvcir.2019.102602).
- [11] ZUO Wangmeng, WU Xiaohe, LIN Liang, *et al.* Learning support correlation filters for visual tracking[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019, 41(5): 1158-1172. doi: [10.1109/TPAMI.2018.2829180](https://doi.org/10.1109/TPAMI.2018.2829180).
- [12] ZABIH R and WOODFILL J. Non-parametric local transforms for computing visual correspondence[C]. The Third European Conference on Computer Vision, Stockholm, Sweden, 1994: 151-158.
- [13] WU Yi, LIM J, and YANG M H. Object tracking benchmark[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(9): 1834-1848. doi: [10.1109/TPAMI.2014.2388226](https://doi.org/10.1109/TPAMI.2014.2388226).
- [14] KRISTAN M, BERG J, LEONARDIS A, *et al.* The seventh visual object tracking VOT2019 challenge results[C]. 2019 IEEE/CVF International Conference on Computer Vision Workshop, Seoul, South Korea, 2019: 2206-2241.
- [15] IRANMEHR A, MASNADI-SHIRAZI H, and VASCONCELOS N. Cost-sensitive support vector machines[J]. *Neurocomputing*, 2019, 343: 50-64. doi: [10.1016/j.neucom.2018.11.099](https://doi.org/10.1016/j.neucom.2018.11.099).
- [16] RAMANAN D. Dual coordinate solvers for large-scale structural SVMs[J]. arXiv: 1312.1743v2, 2014.
- [17] DANELLJAN M, HÄGER G, KHAN F S, *et al.* Convolutional features for correlation filter based visual tracking[C]. 2015 IEEE International Conference on Computer Vision Workshop, Santiago, Chile, 2015: 621-629.
- [18] WANG Qiang, ZHANG Li, BERTINETTO L, *et al.* Fast online object tracking and segmentation: A unifying approach[C]. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, USA, 2019: 1328-1338.
- [19] LI Peixia, CHEN Boyu, OUYANG Wanli, *et al.* GradNet: Gradient-guided network for visual object tracking[C]. 2019 IEEE/CVF International Conference on Computer Vision, Seoul, South Korea, 2019: 6161-6170.
- [20] LI Xin, MA Chao, WU Baoyuan, *et al.* Target-aware deep tracking[C]. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, USA, 2019: 1369-1378.
- [21] BERTINETTO L, VALMADRE J, GOLODETZ S, *et al.* Staple: Complementary learners for real-time tracking[C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, USA, 2016: 1401-1409.
- 袁广林: 男, 1973年生, 博士, 副教授, 主要研究方向为图像处理、计算机视觉、机器学习及其应用等。
- 孙子文: 男, 1996年生, 硕士生, 研究方向为计算机视觉、机器学习。
- 秦晓燕: 女, 1980年生, 硕士, 讲师, 主要研究方向为计算机视觉、机器学习等。
- 夏良: 男, 1980年生, 硕士, 副教授, 主要研究方向为计算机视觉、机器学习、大数据等。
- 朱虹: 女, 1987年生, 硕士, 讲师, 主要研究方向为计算机视觉、图像处理等。

责任编辑: 马秀强