

文章编号 1004-924X(2016)08-2037-13

## 采用核相关滤波器的长期目标跟踪

杨德东, 蔡玉柱\*, 毛宁, 杨福才

(河北工业大学 控制科学与工程学院, 天津 300130)

**摘要:**针对核相关滤波器(KCF)跟踪算法在目标跟踪中存在尺度变化、严重遮挡、相似目标干扰和出视野情况下跟踪失败等问题,提出了一种基于KCF的长期目标跟踪算法。该算法在分类器学习中加入空间正则化,利用基于样本区域空间位置信息的空间权重函数调节分类器系数,使分类器学习到更多负样本和未破坏的正样本,从而增强学习模型的判别力。然后,在检测区域利用Newton方法完成迭代处理,求取分类器最大响应位置及其目标尺度信息。最后,对最大响应位置的目标进行置信度比较,训练在线支持向量机(SVM)分类器,以便在跟踪失败的情况下,重新检测到目标而实现长期跟踪。采用OTB-2013评估基准50组视频序列验证了本文算法的有效性,并与30种其他跟踪方法进行了对比。结果表明:本文提出的算法跟踪精度为0.813,成功率为0.629,排名第一,相比传统KCF算法分别提高了9.86%和22.3%。在目标发生显著尺度变化、严重遮挡、相似目标干扰和出视野等复杂情况下,本文方法均具有较强的鲁棒性。

**关键词:**核相关滤波器;长期目标跟踪;空间正则化;支持向量机(SVM);在线SVM分类器

**中图分类号:**TP391 **文献标识码:**A **doi:**10.3788/OPE.20162408.2037

## Long-term object tracking based on kernelized correlation filters

YANG De-dong, CAI Yu-zhu\*, MAO Ning, YANG Fu-cai

(College of Control Science and Engineering, Hebei University of Technology, Tianjin 300130, China)

\* Corresponding author, E-mail: caiyuzhu001@sina.com

**Abstract:** As Kernelized Correlation Filters (KCF) tracking algorithm has poor performance in handling scale-variant, heavy occlusion, similar target interfere and out of view, this paper proposes a long-term tracking approach based on the KCF. Firstly, a spatial regularization component was introduced in the learning of a classifier, the classifier coefficients were penalized depending on the weight function of spatial location information in sample locations. By which the classifier could learn significantly larger set of negative training samples and uncorrupted positive samples, so that the discriminative power of learned model was increased. Then, the Newton method was used to complete the iteration and obtain the maximizing response location and target score of the classifier in the detection area. Finally, to re-detect the target in the case of tracking failure and achieve a long-term tracking, the confidence of the target with the maximum response score was compared and an online Support Vector Machine (SVM) classifier was trained. To verify the feasibility of the proposed algorithm, fifty groups of OTB-2013 benchmark video sequences were tested and the obtained results were compared with thirty kinds of other tracking algorithms. Experimental results indicate that the precision and

收稿日期:2016-04-18;修订日期:2016-06-22.

基金项目:国家自然科学基金资助项目(No. 61203076);天津市自然科学基金资助项目(No. 13JCQNJC03500)

success rate from the proposed method are respectively 0.813 and 0.629, ranking first. Compared with traditional KCF tracking algorithm, the proposed approach respectively improves by 9.86% and 22.3% in the precision and the success rate. Moreover, it is robust to significant scale changing, heavy occlusion, interfere with similar target, out of view and other complex scenes.

**Key words:** kernelized correlation filter; long-term object tracking; spatial regularization; Support Vector Machine (SVM); online SVM classifier

## 1 引言

目标跟踪是计算机视觉的最重要组成部分之一,其可用于监控、人机交互和医疗图像<sup>[1]</sup>。目标跟踪,是在视频序列的初始帧,给定目标的初始状态(比如位置和尺寸),并在随后的视频序列中估计出目标的状态。影响跟踪算法性能的因素很多,比如尺度变化、遮挡、出视野等,目前也不存在可以处理所有应用场景的跟踪方法。

近年来,对目标跟踪问题的研究取得了许多进展,涌现了很多跟踪算法。根据目标外观模型的表达策略,可将跟踪方法分为生成式方法<sup>[2-3]</sup>和判别式方法<sup>[4-12]</sup>。生成式方法,学习一个外观模型来表示目标,然后在图像区域进行搜索,将与外观模型最相似的区域作为目标。生成式方法丢弃了在目标周围能够将目标从背景中区分出来的有用信息。判别式方法是将跟踪看成一个二元分类的问题,其利用训练数据学习到一个分类器,将目标从背景中分离出来。

最近,已有学者将基于相关滤波器的方法<sup>[13-20]</sup>成功应用在目标跟踪问题中。Bolme等<sup>[13]</sup>提出了一种在灰度图像上学习一个最小输出平方误差和(MOSSE)相关滤波器的跟踪算法,该算法只需要一个目标区域的样本图像训练目标外观模型,通过离散傅里叶变换将目标与所有候选区域之间的相似度计算转换到频域,显著地提升了跟踪算法的运行速度。Henriques等<sup>[14]</sup>提出了循环结构的检测跟踪(CSK)算法,该算法对训练样本进行循环移位,可近似看成对目标的密集采样,从而得到大量训练样本的训练分类器。另外,对候选图像块同样采用循环移位,构造大量候选图像块区域用于分类器的检测过程。分类器的训练和检测过程都可以利用离散傅里叶变换在频域实现快速计算。Danelljan等<sup>[15]</sup>提出了自适应

颜色属性视觉跟踪(CN)算法,使用颜色属性扩展CSK跟踪器,并且采用自适应降维技术将11维的颜色属性特征向量降到了2维,实时地选择比较明显的颜色属性特征,实现了自适应颜色跟踪。Zhang等<sup>[16]</sup>提出时空上下文跟踪(STC)算法,主要利用深度时空上下文信息,将目标周围的背景信息加入到卷积滤波器模型的训练当中,以减弱部分遮挡对目标的影响,对遮挡因素具有一定的抵抗力。Henriques等<sup>[17]</sup>在CSK的基础上提出核相关滤波器(KCF)跟踪算法,使用方向梯度直方图(HoG)<sup>[21]</sup>特征代替原来的灰度值特征,将相关滤波器由单通道扩展到多通道,提升了跟踪性能。

为解决相关滤波器方法无法处理目标尺度变化问题,Li等<sup>[18]</sup>提出了一种基于特征融合的尺度自适应核相关滤波器(SAMF)算法,其融合了颜色属性特征和HoG特征,应用尺度池解决尺度变化;文献<sup>[19]</sup>在传统的基于检测的目标跟踪框架下,设计了一种尺度估计策略,实现了自适应尺度目标跟踪;文献<sup>[20]</sup>提出了一种基于相关滤波器和尺度金字塔的尺度估计机制,并设计了遮挡检测模块,降低遮挡对尺度变化估计的影响。

当目标受尺度变化、严重遮挡、相似目标干扰和出视野等因素影响时,KCF算法很容易出现跟踪失败的问题。针对这一问题,本文在KCF的基础上,引入空间正则化和在线SVM分类器重定位组件,解决了KCF跟踪器在跟踪过程中目标因严重遮挡、相似目标干扰和出视野因素而造成的跟踪失败问题,实现了基于KCF的长期跟踪。为验证本文方法的有效性,利用OTB-2013<sup>[22]</sup>评估基准的50组完全标注的视频序列以及跟踪-学习-检测(TLD)<sup>[6]</sup>算法提供的两组长期跟踪视频序列进行测试,并与KCF及OTB-2013评估基准所包含的29种跟踪算法进行了对比。

## 2 KCF跟踪器

KCF<sup>[17]</sup>跟踪器是一种基于检测的密集采样的跟踪器,利用循环矩阵理论在目标区域进行移位,为分类器构造了大量样本,解决了稀疏采样的样本冗余性问题,提升了跟踪性能和跟踪速度,并在OTB-2013评估基准中取得了最好的跟踪性能和高速的运行速度。

### 2.1 循环矩阵

针对传统的基于检测的跟踪算法稀疏采样的样本存在冗余性问题,KCF跟踪器利用循环矩阵理论在目标区域进行循环移位,为分类器构造大量的训练样本。KCF跟踪器利用基准样本(正样本)和通过循环移位获得虚拟样本(负样本)训练分类器。为表述简单,记基准样本 $\mathbf{x}$ 是 $n \times 1$ 向量。利用置换矩阵 $\mathbf{P}$ 对基准样本 $\mathbf{x}$ 做循环移位操作,移位后的样本为 $\mathbf{P}^l \mathbf{x}$ ,故获取训练分类器的样本集合 $\{\mathbf{P}^l \mathbf{x} | l=0,1,\dots,n-1\}$ 。其中置换矩阵

$$\mathbf{P} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}, \quad (1)$$

循环矩阵 $\mathbf{X}$ ,将对 $n \times 1$ 向量 $\mathbf{x}$ 所有可能的移位进行组合:

$$\mathbf{X} = \mathbf{C}(\mathbf{x}) = \begin{bmatrix} x_1 & x_2 & x_3 & \cdots & x_n \\ x_n & x_1 & x_2 & \cdots & x_{n-1} \\ x_{n-1} & x_n & x_1 & \cdots & x_{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_2 & x_3 & x_4 & \cdots & x_1 \end{bmatrix}, \quad (2)$$

其中循环矩阵 $\mathbf{X}$ 的第一行为向量 $\mathbf{x}$ ,第二行是将向量 $\mathbf{x}$ 元素向右移动一位,其他行以此类推。由于循环特性,每移位 $n$ 次,可以周期性获得基准样本。同样,可以等效地看成循环矩阵 $\mathbf{X}$ 的前半部分对向量 $\mathbf{x}$ 元素的移位沿着正方向,而循环矩阵 $\mathbf{X}$ 的后半部分的移位则沿着反方向。

循环矩阵 $\mathbf{X}$ 可通过离散傅里叶变换变成对角矩阵,可以表示为

$$\mathbf{X} = \mathbf{F} \text{diag}(\hat{\mathbf{x}}) \mathbf{F}^H, \quad (3)$$

其中 $\mathbf{F}$ 为离散傅里叶变换矩阵,符号记为离散傅里叶变换, $\hat{\mathbf{x}}$ 为向量 $\mathbf{x}$ 的离散傅里叶变换, $\mathbf{H}$ 表示矩阵共轭转置。

### 2.2 分类器训练

在许多实际问题中,由于岭回归对于任意输入都有简单的闭式最优解,不需要复杂的迭代、最小化序列优化等,不仅降低了计算复杂度,还能够获得与支持向量机(SVM)相近的分类性能。

为此,KCF聚焦岭回归,利用一组训练样本 $\{\mathbf{x}_i, y_i\}$ ,从中寻找对样本 $\mathbf{x}_i$ 和回归目标 $y_i$ 最小化平方误差的函数 $f(\mathbf{z}) = \mathbf{w}^T \mathbf{z}$ ,其中 $\mathbf{w}$ 为分类器参数, $\mathbf{z}$ 为检测图像块

$$\min_{\mathbf{w}} \sum_i (f(\mathbf{x}_i) - y_i)^2 + \lambda \|\mathbf{w}\|^2, \quad (4)$$

其中 $\lambda$ 是控制过拟合的正则化项参数。

由Representer定理<sup>[23]</sup>给出公式(4)最小化解:对输入的训练样本线性组合

$$\mathbf{w} = \sum_i \alpha_i \varphi(\mathbf{x}_i), \quad (5)$$

其中 $\varphi(\mathbf{x}_i)$ 为将训练样本映射到高维特征空间的函数。

定义核函数 $k(\mathbf{x}, \mathbf{x}') = \varphi^T(\mathbf{x}) \varphi(\mathbf{x}')$ ,其中核函数 $k$ 为高斯核函数或者多项式核函数。 $n \times n$ 核矩阵 $\mathbf{K}$ 由所有样本对的点乘 $K_{ij}$ 组成

$$K_{ij} = k(\mathbf{x}_i, \mathbf{x}_j) \quad (6)$$

核函数 $k$ 是酉不变核函数,则以 $K_{ij}$ 为元素的 $\mathbf{K}$ 核矩阵是循环矩阵<sup>[14]</sup>。其中,径向基核函数、多项式核函数和高斯核函数都是酉不变核函数。

然后,基于核的正则化最小二乘<sup>[24]</sup>给出岭回归简单的闭式解

$$\boldsymbol{\alpha} = (\mathbf{K} + \lambda \mathbf{I})^{-1} \mathbf{y}, \quad (7)$$

其中 $\mathbf{K}$ 为核矩阵, $\mathbf{I}$ 为单位矩阵,向量 $\mathbf{y}$ 的元素为 $y_i$ ,向量 $\boldsymbol{\alpha}$ 的元素为 $\alpha_i$ 。

由于核矩阵 $\mathbf{K}$ 具有循环特性,利用公式(3)的性质对公式(7)进行离散傅里叶变换:

$$\hat{\boldsymbol{\alpha}} = \frac{\hat{\mathbf{y}}}{\hat{\mathbf{k}}^{xx} + \lambda}, \quad (8)$$

其中 $\hat{\mathbf{k}}^{xx}$ 为核矩阵 $\mathbf{K} = \mathbf{C}(\mathbf{k}^{xx})$ 的第一行元素。

### 2.3 快速检测

KCF跟踪算法,对于新输入图像块 $\mathbf{z}$ ,分类器的响应为:

$$f(\mathbf{z}) = \mathbf{w}^T \mathbf{z} = \sum_i \alpha_i k(\mathbf{z}, \mathbf{x}_i). \quad (9)$$

对图像块 $\mathbf{z}$ 进行循环移位构造候选图像块,可利用候选图像块的循环结构,同时计算出分类器对检测区域中的所有位置的响应。

记 $\mathbf{K}^c$ 为训练样本和所有候选图像块的核矩

阵。由于训练样本由基准样本  $\mathbf{x}$  循环移位构成, 候选图像块由基准图像块  $\mathbf{z}$  循环移位构成, 利用酉不变核函数定理, 可知  $\mathbf{K}^z$  为循环矩阵。

$$\mathbf{K}^z = C(\mathbf{k}^{xz}), \quad (10)$$

$\mathbf{K}^z$  表示训练样本和所有候选图像块之间的核矩阵,  $\mathbf{k}^{xz}$  表示  $\mathbf{x}$  和  $\mathbf{z}$  之间的核相关。

利用公式(9)和公式(10), 可计算出所有候选图像块的分器响应

$$\mathbf{f}(\mathbf{z}) = (\mathbf{K}^z)^T \boldsymbol{\alpha}, \quad (11)$$

其中  $\mathbf{f}(\mathbf{z})$  是一个向量, 包含  $\mathbf{z}$  所有循环移位的输出, 即所有检测结果。

为有效计算公式(11), 可利用  $\mathbf{K}^z$  的循环特性, 及公式(3)的性质对公式(11)进行离散傅里叶变换:

$$\hat{\mathbf{f}}(\mathbf{z}) = \hat{\mathbf{k}}^{xz} \cdot \hat{\boldsymbol{\alpha}}, \quad (12)$$

其中  $\cdot$  表示点乘。

### 3 采用 KCF 的长期跟踪算法

由 2.1 节可知, KCF 跟踪算法中, 训练样本周期性循环移位产生。然而, 这种周期性循环移位对样本产生了不必要的边界效应。

KCF 算法在理论上对跟踪性能的影响主要体现在以下几方面。(a) 由于 KCF 算法限制了分类器的检测区域和训练区域为目标大小的四倍, 无法跟踪快速运动目标; 负样本数量比较少, KCF 算法的学习模型过拟合, 相似目标容易混淆; (b) 由于周期性循环移位对样本产生了不必要的边界效应, 限制了 KCF 算法只在靠近搜索区域的中心位置跟踪准确, 因而在严重遮挡和出视野情况下表现不好。

本文通过扩大样本区域, 引入空间正则化和在线 SVM 分类器重定位组件, 解决了 KCF 算法在尺度变化、严重遮挡、相似目标干扰和出视野因素下跟踪失败的问题。

#### 3.1 空间正则化组件

空间正则化, 是在分类器学习过程中利用正则化权重  $\tau$  调节分类器参数  $\mathbf{w}$ 。正则化权重  $\tau$  是基于分类器搜索范围的先验信息分布的。在背景区域, 赋予较高的正则化权重  $\tau$  来调节分类器参数  $\mathbf{w}$ , 以便减轻背景信息对分类器在学习过程中的影响。

假定所有样本具有同样的空间尺寸  $M \times N$ ,

即样本的空间位置  $(m, n) \in \Omega; = \{0, \dots, M-1\} \times \{0, \dots, N-1\}$ 。在学习过程中, 为调节分类器参数  $\mathbf{w}$ , 引入空间权重函数  $\tau: \Omega \rightarrow \mathcal{R}$ 。空间正则化权重根据空间位置决定分类器参数  $\mathbf{w}$  的重要性, 背景区域赋予较高空间权重, 目标区域内赋予较小的空间权重。利用空间正则化权重  $\tau$  代替公式(2)中的正则化项  $\lambda$ :

$$\min_{\mathbf{w}} \sum_i (f(\mathbf{x}_i) - y_i)^2 + \|\tau \cdot \mathbf{w}\|^2, \quad (13)$$

其中空间正则化权重函数

$$\tau(m, n) = u + \eta \{(m/P)^2 + (n/Q)^2\}. \quad (14)$$

最小空间正则化权重  $\tau = u = 0.1$ , 正则化影响因数  $\eta = 3$ ,  $P \times Q$  表示目标尺寸。

由帕塞瓦尔定理可知, 函数平方的和等于其离散傅里叶变换式平方之和。对公式(13)应用帕塞瓦尔定理

$$\min_{\hat{\mathbf{w}}} \sum_i (\hat{x}_i \cdot \hat{\mathbf{w}} - \hat{y}_i)^2 + \left\| \frac{\hat{\tau}}{MN} * \hat{\mathbf{w}} \right\|^2, \quad (15)$$

其中  $*$  表示循环卷积。

公式(15)第二项遵循逆离散傅里叶变换的卷积属性, 公式(15)向量化:

$$\min_{\hat{\mathbf{w}}} \sum_i \|D(\hat{x}_i) \hat{\mathbf{w}} - \hat{y}_i\|^2 + \left\| \frac{c(\hat{\tau})}{MN} \hat{\mathbf{w}} \right\|^2, \quad (16)$$

其中, 粗体表示相应标量函数的向量化,  $D(\mathbf{v})$  表示向量  $\mathbf{v}$  的元素在其对角线上的对角矩阵。大小为  $MN \times MN$  矩阵  $c(\hat{\tau})$  表示对函数  $\hat{\tau}$  的二维循环卷积, 即  $c(\hat{\tau}) \hat{\mathbf{w}} = \text{vec}(\hat{\tau} * \hat{\mathbf{w}})$ 。矩阵  $c(\hat{\tau})$  每行元素为  $\hat{\tau}$  元素的周期性循环移位。

实值函数的离散傅里叶变换是埃尔米特对称。因此, 通过设置埃尔米特对称离散傅里叶参数  $\hat{\mathbf{w}}$  最小化公式(15), 等效于设置  $\omega$  中的实值部分最小化公式(13)。通过设计一个  $MN \times MN$  酉矩阵, 每行至多两个非零元素,  $\tilde{\mathbf{w}} = B\hat{\mathbf{w}}$ , 获取  $\hat{\mathbf{w}}$  中的实值部分  $\tilde{\mathbf{w}}$ 。

本文定义  $\tilde{\mathbf{y}}_i = B\hat{y}_i$ ,  $\mathbf{D}_i = BD(\hat{x}_i)B^H$ ,  $\mathbf{C} = \frac{1}{MN} \cdot Bc(\hat{\tau})B^H$ , 其中  $H$  表示矩阵的共轭转置。公式(16)可等效表达为

$$\min_{\tilde{\mathbf{w}}} \sum_i \|\mathbf{D}_i \tilde{\mathbf{w}} - \tilde{\mathbf{y}}_i\|^2 + \|\mathbf{C} \tilde{\mathbf{w}}\|^2, \quad (17)$$

定义  $\boldsymbol{\tau}'$  为大小为  $MN \times MN$  块对角矩阵, 每个对角块等于  $\mathbf{C}$ 。

$$\min_{\tilde{\mathbf{w}}} \sum_i \|\mathbf{D}_i \tilde{\mathbf{w}} - \tilde{\mathbf{y}}_i\|^2 + \|\boldsymbol{\tau}' \tilde{\mathbf{w}}\|^2, \quad (18)$$

最后利用求解正规方程  $\mathbf{A}_t \tilde{\mathbf{w}} = \tilde{\mathbf{b}}_t$ , 最小化公式(18)

$$\mathbf{A}_t = \sum_i \mathbf{D}_i^T \mathbf{D}_i + \boldsymbol{\tau}'^T \boldsymbol{\tau}', \quad (19)$$

$$\tilde{\mathbf{b}}_t = \sum_i \mathbf{D}_i^T \tilde{\mathbf{y}}_i. \quad (20)$$

### 3.2 模型训练

在训练阶段,首先提取一个以目标位置为中心的样本  $\mathbf{x}_t$ 。其中,  $t$  表示当前帧。然后,利用学习效率  $\gamma \geq 0$ , 采用线性插值对当前帧参数与之前帧的相关参数进行整合。更新公式(19)中的  $\mathbf{A}_t$  和公式(20)中的  $\tilde{\mathbf{b}}_t$

$$\mathbf{A}_t = (1 - \gamma) \mathbf{A}_{t-1} + \gamma (\mathbf{D}_t^T \mathbf{D}_t + \boldsymbol{\tau}'^T \boldsymbol{\tau}'), \quad (21)$$

$$\tilde{\mathbf{b}}_t = (1 - \gamma) \tilde{\mathbf{b}}_{t-1} + \gamma \mathbf{D}_t^T \tilde{\mathbf{y}}_t. \quad (22)$$

第一帧, 设定  $\mathbf{A}_1 = \mathbf{D}_1^T \mathbf{D}_1 + \boldsymbol{\tau}'^T \boldsymbol{\tau}'$ ,  $\tilde{\mathbf{b}}_1 = \mathbf{D}_1^T \tilde{\mathbf{y}}_1$ 。针对整个视频序列, 提前计算出正则化矩阵  $\boldsymbol{\tau}'^T \boldsymbol{\tau}'$ , 并使其保持不变。

### 3.3 快速子网格检测

在检测阶段, 利用在之前帧更新的分类器  $\hat{\mathbf{w}}_{t-1}$  估计第  $t$  帧的目标位置。本文利用插值方法, 实现密集像素检测。记  $\hat{s} = F(f(\mathbf{z})) = \hat{\mathbf{w}} \cdot \hat{\mathbf{z}}$ , 为图像块  $\mathbf{z}$  的检测值  $f(\mathbf{z})$  的离散傅里叶变换。检测区域的连续位置  $(u, v) \in [0, M] \times [0, N]$  的检测值, 可以插值表示为:

$$s(u, v) = \frac{1}{MN} \sum_0^{M-1} \sum_0^{N-1} \hat{s}(m, n) e^{i2\pi(\frac{m}{M}u + \frac{n}{N}v)}, \quad (23)$$

其中  $i$  表示虚部。快速子网络检测, 获取的是检测值最大化的子网格位置  $(u^*, v^*)$ :

$$(u^*, v^*) = \operatorname{argmax}_{(u, v) \in [0, M] \times [0, N]} s(u, v), \quad (24)$$

利用  $s = F^{-1}(\hat{\mathbf{w}} \cdot \hat{\mathbf{z}})$ , 求取分类器第一次在所有网格位置的检测值  $s(m, n)$ , 并从中选取分类器最大检测值的子网格位置  $(u^{(0)}, v^{(0)})$ , 作为初始位置估计。然后, 以子网格位置  $(u^{(0)}, v^{(0)})$  为开始位置, 利用 Newton 方法迭代处理公式(24), 经过几次迭代后, 便可获取分类器最大检测值的子网格位置。

类似于参考文献[18], 采用尺度池解决目标尺度变化的问题。以前一帧的目标位置为中心, 利用相对于当前目标尺度的尺度  $a^r$  提取样本  $\{z_r\}$ , 其中  $r \in \left\{ \left\lfloor \frac{1-S}{2} \right\rfloor, \dots, \left\lfloor \frac{S-1}{2} \right\rfloor \right\}$ , 其中:  $S$  表示尺度数量,  $a$  表示尺度增加因子。在计算特征

之前, 根据尺度  $a^r$  调整图片大小, 构造样本  $z_r$ 。

本文利用子网格插值策略, 获取分类器对样本  $z_r$  分类器得分  $s_r$  的最大值, 利用最大分类器得分  $s_r$  的子网格位置更新目标位置。对样本  $z_r$  的每个尺度层, 做独立的快速子网格检测, 从中找到获取分类器最高得分的尺度层, 用以更新目标尺度。

### 3.4 重定位组件

长期跟踪过程中, 重定位组件对于目标因受到长期遮挡或者出视野等因素导致跟踪失败后重新定位目标位置至关重要。对于重定位组件, 由于跟踪实时性的要求, 本文并不是对每帧图像都训练一个目标相关滤波器, 而是独立地训练一个在线 SVM 分类器。

长期跟踪算法 TLD 采用对视频序列的每帧图像都执行重检测的策略。而本文采用阈值  $\theta_{re}$  激活策略, 激活重定位组件。当  $\max(s_r) < \theta_{re}$  时, 即当分类器检测值最大值小于重定位激活阈值时, 激活目标重检测, 利用在线 SVM 分类器扫描整帧图像, 获取在线 SVM 分类器的重检测结果, 通过比较目标的置信度确定是否采用重检测结果。

本文采用一种半监督学习方法训练在线 SVM 分类器, 以增强重定位组件的定位性能。利用标记集  $L_l = \{[x_{re}, y_{re}]\}$  和未标记集  $U_u$ , 其中  $l \ll u$ ,  $x_{re}$  和  $y_{re}$  分别为重定位组件训练在线 SVM 分类器的样本及标签。标记集选取第一帧给定目标图像块和与第一帧给定目标框重叠率在  $(0.6, 1)$  之间的图像块作为正样本, 而负样本为与第一帧给定目标框重叠率在  $(0, 0.2)$  之间的图像块。与正负样本来源于第一帧的标记集不同, 未标记集在跟踪过程(帧数  $i \geq 2$ ) 中产生, 选取与跟踪的目标框重叠率在  $(0.6, 1)$  之间的图像块作为正样本, 而负样本为与跟踪的目标框重叠率在  $(0, 0.2)$  之间的图像块。

首先, 将标签集  $L_l$  加入到训练集中, 初始化训练在线 SVM 分类器, 并获取在线 SVM 分类器的初始化参数  $w_{re}^0$ ; 然后, 在未标记集  $U_u$  上, 将初始训练的在线 SVM 分类器经  $k$  次迭代获取在线分类器参数  $w_{re}^k$ ,  $\forall x_{re(\omega)} \in U_u, y_{re(\omega)}^k = w_{re}^k \cdot x_{re(\omega)} + b$ 。利用 P-N 学习的结构化约束[6], 识别在线 SVM 分类器赋予未标记数据的标签是否与结构化约束相矛盾。只有当  $\max(s_r) > \theta_u$  时, 分类器

检测值最大值大于目标外观阈值,从目标周围选取正负样本,加入训练集并更新,训练更新 SVM 分类器。

## 4 实验结果与分析

为验证本文算法的有效性,利用 OTB-2013<sup>[22]</sup>中大约 29 000 帧图像的 50 个完全标注的视频数据集包含 11 种属性进行评估,并与 KCF 以及 OTB-2013 代码库中所包含的 29 种跟踪算法进行了对比。OTB-2013 代码库所包含了以下跟踪算法:CSK, Struck, SCM, TLD, CT, VTD, VTS, CXT, IVT, ASLA, DFT, LSK, CPF, LOT, VR-V, KMS, L1APG, MTT, MIL, OAB, SemiT, ORIA, SMS, Frag, BSBT。OTB-2013 视频数据集包含了以下属性:光照变化(IV)、尺度变化(SV)、遮挡(OCC)、形变(DEF)、运动模糊(MB)、快速运动(FM)、平面内旋转(IPR)、平面外旋转(OPR)、出视野(OV)、复杂背景(BC)、低分辨率(LR)。

### 4.1 实验环境及参数设置

本实验基本配置,Matlab2013a, Intel (R) Core(TM) i3-2350 CPU,主频 2.3 GHz,4 GB 内存。类似于 KCF 算法,本文算法采用 HoG 特征,大小为 4 pixel×4 pixel 的单元。采用 M×N 的网格单元表示样本大小,并与目标框区域成比例。KCF 算法将样本区域设置为初始目标框区域的 4 倍,本文充分考虑到快速运动目标连续两帧之间真实目标框的中心位置的偏移大于 20 pixel 这一情况,以及充分考虑分类器训练、检测效率,将样本区域设置为初始目标框区域的 16 倍。空间正则化权重函数为  $\tau(m, n)$ ,最小空间正则化权重  $\tau = u = 0.1$ ,正则化影响因数  $\eta = 3$ , $P \times Q$  表示目标尺寸,尺度数量  $S = 7$ ,尺度增量因子  $a = 1.01$ 。重定位组件设置如下,设置  $\theta_r = 0.5$  用以激活在线 SVM 分类器,设置  $\theta_a = 0.5$  用以更新重定位目标模型。模型更新中,学习效率  $\gamma = 0.01$ 。对于实验中所有的视频数据集,实验参数均保持一致。

### 4.2 性能评估方法

为了评估跟踪算法的性能,本实验采用了 OTB-2013 中的两种评估方法作为定量评价指标:精确度图(Precision plot)和成功率图(Suc-

cess plot)。

在跟踪精度评估中,广泛应用的是中心位置误差(CLE)。中心位置误差定义为跟踪目标的中心位置和手工标定的准确中心位置之间的平均欧式距离。

精确度图,能够显示目标中心位置误差小于阈值的帧数占视频总帧数的百分比。实验中阈值选定为 20 pixels。

成功率图,能够显示重叠率  $S$  大于给定阈值  $t_0$  的帧数占视频总帧数的百分比。其中,重叠率  $S$  定义为, $S = |R_t \cap R_a| / |R_t \cup R_a|$ , $R_t$  为跟踪的目标框, $R_a$  为准确的目标框, $\cap$  和  $\cup$  分别表示两个区域的交集和并集, $|\cdot|$  为区域内的像素点的数量。成功率给出了阈值  $t_0$  从 0 到 1 变化时,成功的帧数所占的比例,利用成功率图曲线下的面积(AUC)作为跟踪算法性能的评价准则。

### 4.3 性能比较

本实验采用基于精确度图和成功率图的一次通过评估(OPE)方法评估跟踪算法性能。利用 OTB-2013 评估策略,对共计 30 个跟踪算法进行性能比较。为了显示清楚,本文在精确度图和成功率图上只显示排序靠前的 10 个跟踪算法的结果。

#### 4.3.1 定量比较

##### A) 整体性能分析

图 1 为整体性能排名前 10 的 OPE 精确度图和成功率图。依据精确度图和成功率图,两个图中本文算法均排在第一位。在精确度图中,本文算法精确度为 0.813,相比于 KCF 和 CSK 算法分别提高了 9.86% 和 49.1%。在成功率图中,本文算法成功率为 0.629,相比于 KCF 和 CSK 算法分别提高了 22.3% 和 58%。

##### B) 基于数据集属性的性能分析

为充分评估本文算法跟踪性能,利用 OTB-2013 视频数据集的 11 个属性进一步评估本文算法的性能。图 2(彩图见期刊电子版)为排名前 10 的跟踪算法在包含 11 个属性数据集上的精确度图,由图 2 可知,本文算法除了在复杂背景、低分辨率和光照变化这 3 个属性精度图上排名第二,其余 8 个属性的精度图上都排名第一。对比其他算法中效果最好的 KCF 算法,在尺度变化、遮挡、出视野和快速运动属性精度图上,本文算法分别提高了 14.7%、5.74%、23% 和 16.7%。

图 3(彩图见期刊电子版)为排名前 10 的跟

踪算法在包含 11 个属性数据集上的成功率图,本文算法在 11 个属性的成功率图上都排名第一位。在尺度变化、遮挡、出视野和快速运动这 4 个属性的成功率图上,本文算法较 KCF 算法分别提高了 40.5%、20.2%、15.6% 和 23.3%。

对于 OTB-2013 具有遮挡属性的 29 个数据集,本文算法在精确度图和成功率图上均排第一,相比排名第二位的 KCF 分别提高了 5.74% 和 20.2%,取得了不错的改进效果。本文算法的样本区域比 KCF 算法的扩大了 4 倍,并通过加入空间正则化组件,以有效地学习更多的背景信息,获

取更多的负样本和未被破坏的正样本,增强了分类器将目标从背景中区分的能力,即处理遮挡数据集的能力。

对于 OTB-2013 出视野属性的 6 个数据集,本文算法在精确度图和成功率图上均排序第一,相比排名第二位的 KCF 分别提高了 23% 和 15.6%。这是因为本文算法加入了重定位模块,当相关滤波器确定的目标置信度比较低时,采用阈值激活重定位 SVM 分类器方法,重新确定目标位置,并在目标位置区域进行采样训练 SVM 分类器,提高重定位能力。

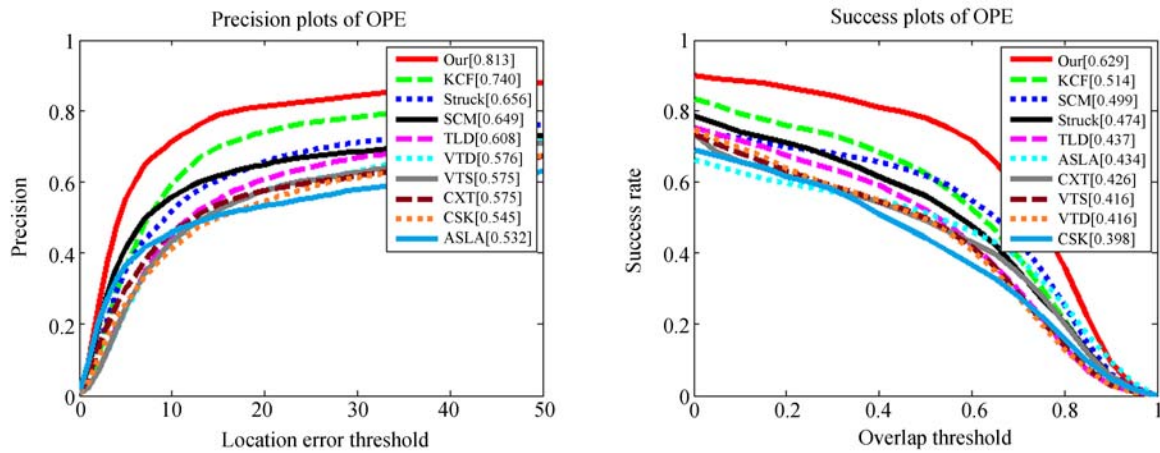
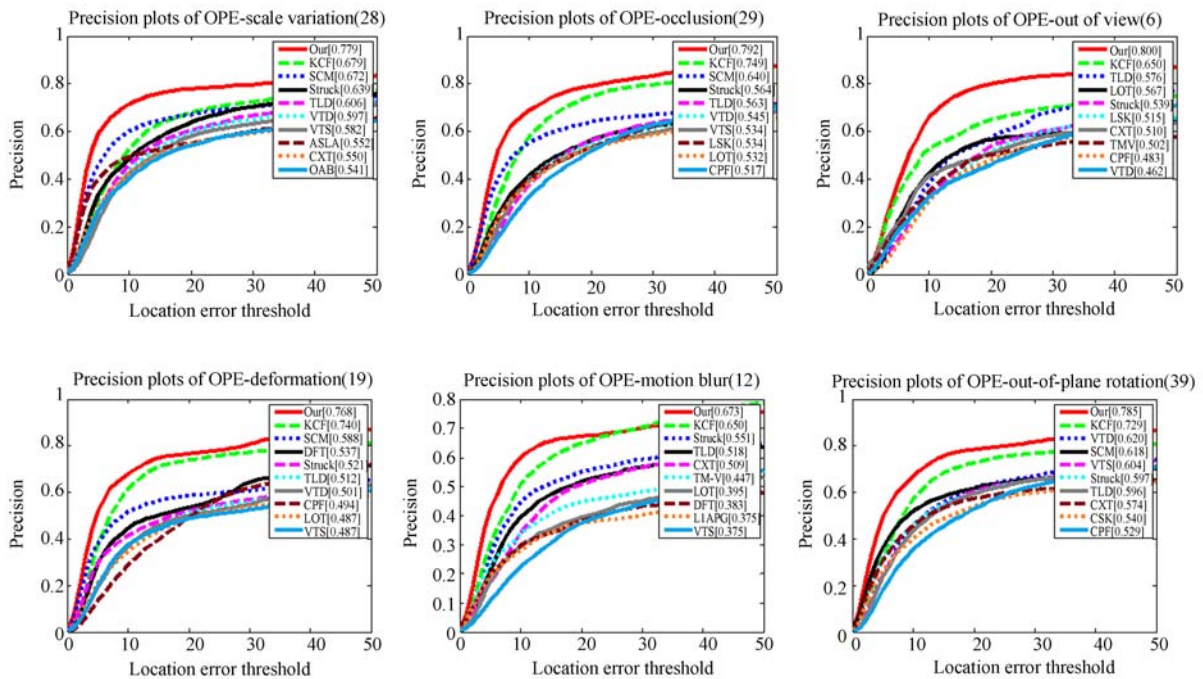


图 1 排名前 10 跟踪算法 OPE 的精确度图和成功率图

Fig. 1 Precision plots and success plots of OPE for the top 10 trackers



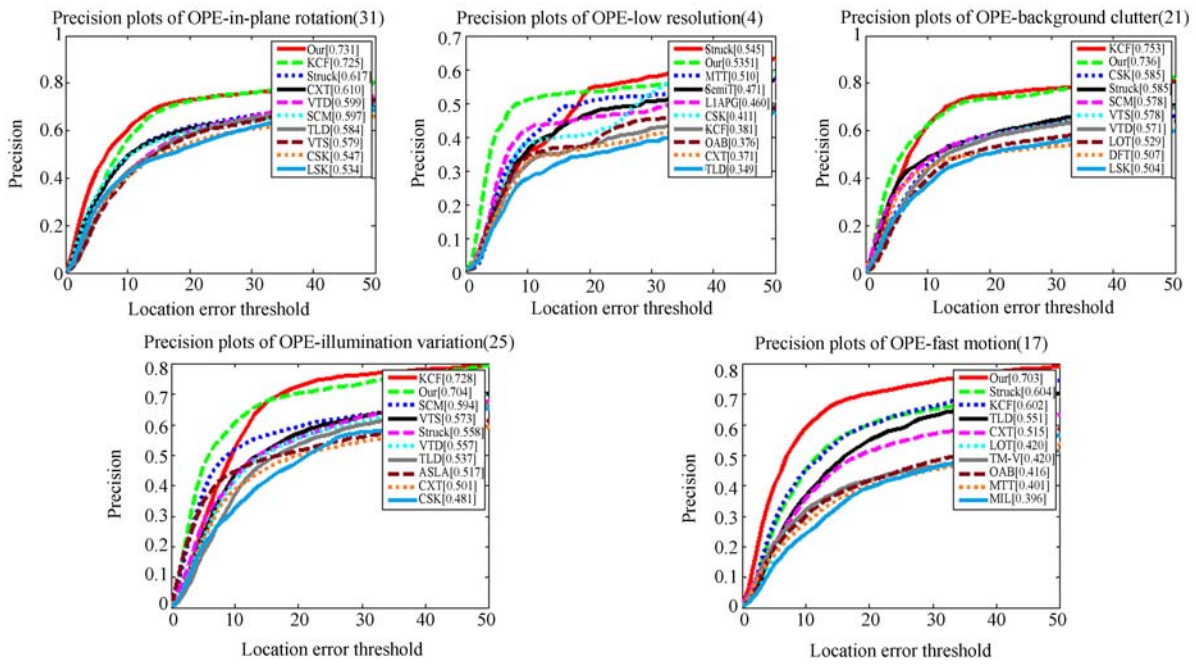
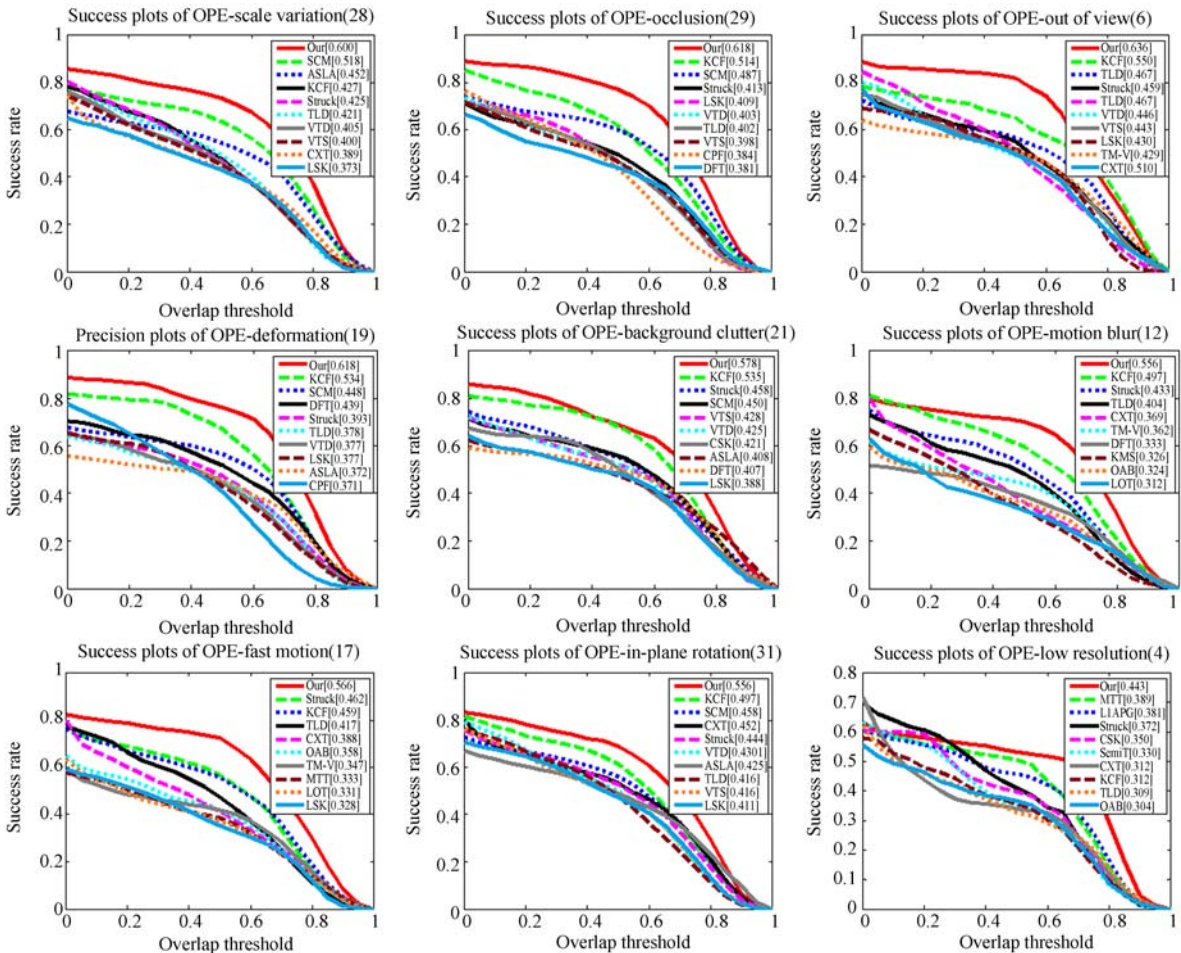


图 2 排名前 10 跟踪算法在 11 个属性的精确度图(标题上的数字表示此种属性的数据集数量)

Fig. 2 Precision plots of videos with 11 different attributes for the top 10 trackers. (The number in the title indicates the number of sequences.)



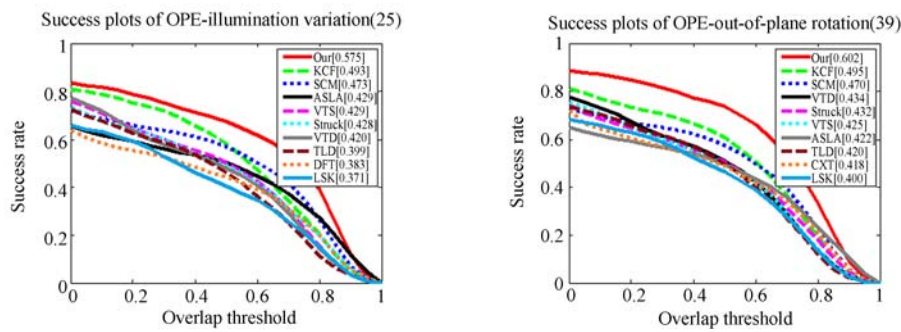


图 3 排名前 10 跟踪算法在 11 个属性的成功率图(标题上的数字表示此种属性的数据集数量)

Fig. 3 Success plots of videos with 11 different attributes for the top 10 trackers. (The number in the title indicates the number of sequences.)

4. 3. 2 定性比较

A) 抗遮挡性能分析

图 4(彩图见期刊电子版)显示了 10 种算法在 3 个经历严重遮挡数据集上的几帧具有代表性跟踪结果。在 Jogging 数据集中,在第 67 帧目标被灯杆几乎完全遮挡。10 种算法中,只有本文算法、TLD 和 CXT 算法能够重新定位到目标,例如第 79 帧、第 130 帧和第 288 帧。在 Woman 数据集中,目标人物经历了长时间持续遮挡,本文算法、SCM 算法和 Struck 算法都能够正确跟踪目

标人物,例如第 165 帧、第 337 帧和第 477 帧。在 Suv 数据集中,目标车辆频繁地被茂密的树木遮挡,例如第 513 帧、第 535 帧和第 778 帧。本文算法、SCM 算法、TLD 算法和 CXT 算法均取得了不错的跟踪效果,例如第 578 帧和 850 帧。在这 3 个经历严重遮挡的数据集中,KCF 算法、CSK 算法都无法准确跟踪目标,表现不好。由以上结果可知,本文算法通过扩大样本区域,利用空间正则化组件对遮挡具有抵抗力。

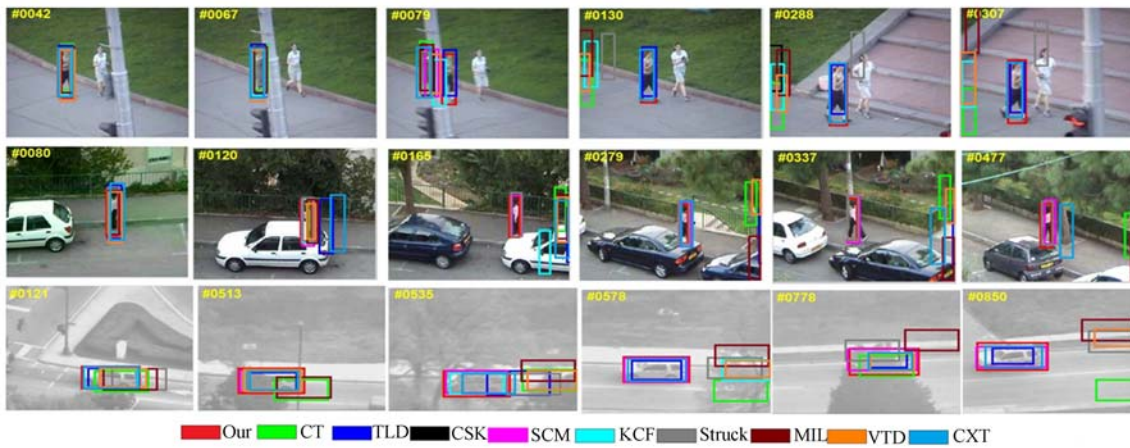


图 4 10 个跟踪算法在经历严重遮挡的数据集 Jogging, Woman, Suv 上定性结果显示

Fig. 4 Qualitative results of 10 trackers over sequences Jogging, Woman and Suv, in which the targets undergo heavy occlusion

B) 相似目标干扰性能分析

图 5(彩图见期刊电子版)显示了 10 种跟踪算法在经历相似目标干扰的数据集上具有代表性的跟踪结果。在数据集 Deer 中,在第 30 帧和第 55 帧目标鹿的头部与周围干扰鹿的头部相似度极高,对目标鹿的跟踪产生一定的干扰。

KCF 算法的跟踪结果(淡青色目标框)与目标鹿头部真实位置有一定的误差,而本文算法的跟踪结果(鲜红色目标框)与目标鹿头部真实位置一致。在数据集 Football 中,由于要跟踪的目标只是橄榄球运动员的头部,而橄榄球运动员都带着同样的头盔,目标运动员与周围的运动员头部

的相似性极高。在第 283 帧,目标运动员头部被 37 号运动员头部遮挡。而在第 310 帧,KCF 算法跟踪到了 37 号运动员的头部,出现跟踪错误;而本文算法能够正确跟踪目标运动员的头部,未被 37 号运动员头部干扰。在数据集 Liquor 中,第

327 帧中左侧的瓶子与中间位置的瓶子具有一定的相似性。在第 550 帧、第 751 帧、第 970 帧和第 1 539 帧,KCF 算法受相似目标的干扰,出现了错误跟踪;而本文算法均能够正确跟踪真实目标,未受相似目标干扰而跟踪失败。

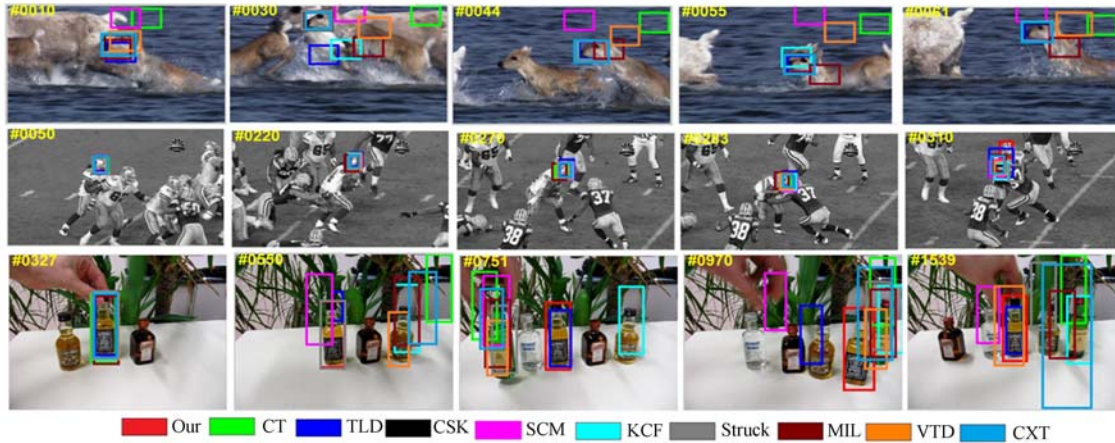


图 5 10 个跟踪算法在经历相似目标干扰的数据集 Deer, Football, Liquor 上定性结果显示

Fig. 5 Qualitative results of the 10 trackers over sequences Deer, Football and Liquor, in which the targets undergo disturbance of similar target.

### C) 尺度变化性能分析

图 6 显示了 10 种跟踪算法在具有较大尺度变化的数据集上具有代表性的跟踪结果。在 Carscale 数据集中,目标车辆经历了显著地尺度变化,例如第 149 帧、第 165 帧、第 185 帧和第 252 帧。CSK 算法由于采用简单的灰度特征进行跟踪,与当目标车辆在第 165 帧受到树木遮挡,导致模型特征发生变化时,出现跟踪失败(如第 185

帧所示)。在这 10 种算法中,本文算法对于 Car-scale 数据集中目标车辆具有最好的跟踪效果,对显著尺度变化具有较强的鲁棒性。在 Car4 数据集中,目标车辆尺度由大变小,并受到严重的光照变化影响,例如第 187 帧和第 303 帧。本文算法、TLD 算法和 SCM 算法均能较好地处理目标车辆的尺度变化,而 CT 算法由于采用随机投影矩阵对高维特征进行降维,故特征选取不显著,在跟踪

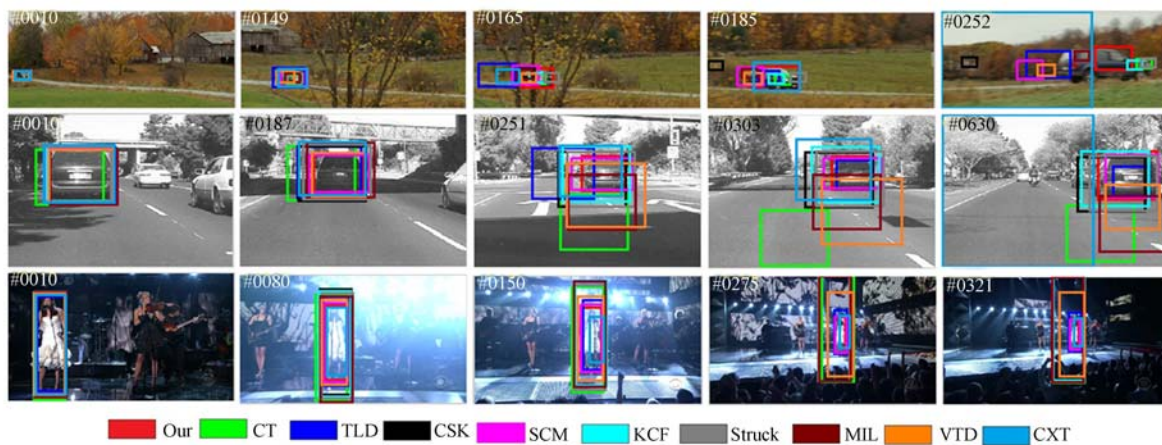


图 6 10 个跟踪算法在经历尺度变化的数据集 CarScale, Car4, Singer1 上定性结果显示

Fig. 6 Qualitative results of the 10 trackers over sequences CarScale, Car4 and Singer1, in which the targets undergo scale variations.

过程中出现跟踪漂移。在 Singer1 数据集中,由于镜头拉远,目标人物尺度由大变小,经历了显著的尺度变化,并且受到严重的光照变化影响,例如第 80 帧和第 150 帧。与 KCF 算法、CSK 算法、CT 算法和 VTD 算法相比,本文算法在处理目标尺度变化问题上具有一定的优越性,对显著的尺度变化具有较强的鲁棒性。

#### D) 出视野性能分析

图 7 显示了 10 种跟踪算法在目标出视野的数据集上具有代表性的跟踪结果。由于超出视野数据集在 OTB-2013 评测数据集中只有 6 个,本文选取了超出视野因素影响比较大的 Lemming 数据集和 TLD 算法所提供的 Motocross 和 Car-chase 数据集进行分析。在 Lemming 数据集中,旅鼠玩偶目标超出视野(如第 556 帧),一段时间

后又完全回到视野中(如第 610 帧),本文算法可以正确跟踪到旅鼠玩偶目标,而 KCF 算法则跟踪失败。在 Lemming 数据集中,多次出现平面外和平面内旋转因素的影响,给跟踪增加了难度,但是本文算法仍然能够正确跟踪。在数据集 Motocross 中,目标经历了超出视野(如第 31 帧和第 493 帧)、外观严重变化(如第 427 帧和第 552 帧)影响,本文算法都能正确跟踪。在 Car-chase 数据集中,目标车辆经历完全出视野一段时间又回到视野中(如第 377 帧和第 389 帧),本文算法和 TLD 算法由于加入了重定位组件,即使在目标完全出视野一段时间后又回到视野中这种复杂情况下仍然能够跟踪正确,而 KCF 算法由于没有重定位模块故在目标出视野一段时间后重回视野后,无法重新定位目标的位置而跟踪失败。

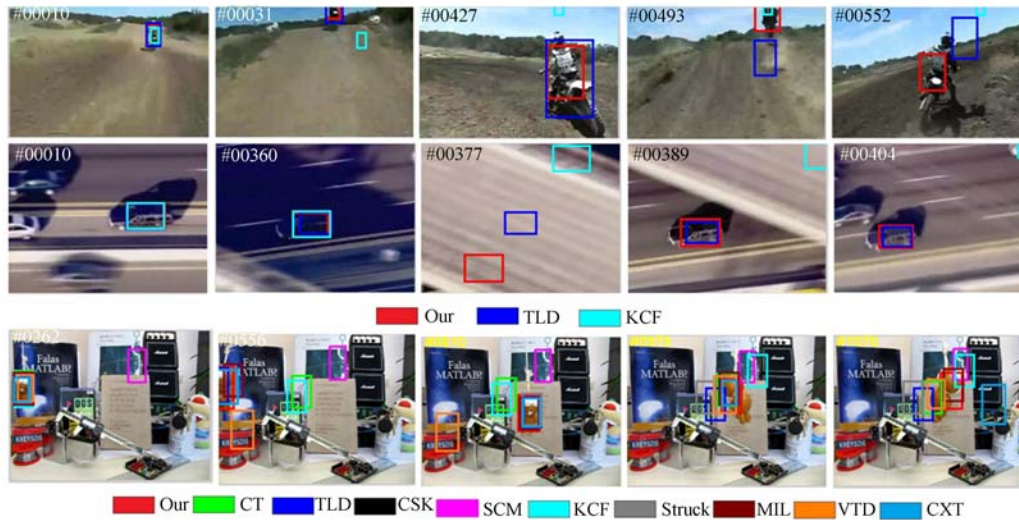


图 7 10 个跟踪算法在经历出视野的数据集 Motocross, Car-chase, Lemming 上定性结果显示

Fig. 7 Qualitative results of the 10 trackers over sequences Motocross, Car-chase and Lemming, in which the targets undergo out of view.

## 5 结 论

针对 KCF 算法在尺度变化、严重遮挡、相似目标干扰、出视野等复杂情况下出现跟踪失败的问题,本文提出了基于核相关滤波器的长期跟踪方法。在模型学习过程中,本文通过扩大样本区域(相比 KCF 算法样本区域扩大了 4 倍),加入空间正则化组件,利用基于样本区域的空间先验信息的空间权重函数,减轻样本区域中的背景信息对于分类器学习的影响。解决了 KCF 算法在产

生训练样本过程中利用周期性假设而带来的边界效应(正样本被损坏),学习更多的未被损坏的正样本和负样本,从而提高了学习模型的判别能力。另外,本文通过加入重定位组件,解决了目标出视野一段时间后重回到视野中跟踪问题。利用 OTB-2013 的 50 个数据集进行试验,结果显示本文方法的整体精度为 0.813,成功率为 0.629,相比 KCF 算法分别提高了 9.86% 和 22.3%。通过基于数据集的属性的定量和定性分析表明,在目标发生显著尺度变化、严重遮挡、相似目标干扰和出视野等复杂场景下,与 OTB-2013 代码库中的

29 种算法以 KCF 算法相比,本文方法具有更强的鲁棒性。

#### 参考文献:

- [1] SMEULDERS A W M, CHU D M, CUCCHIARA R, *et al.*. Visual tracking: An experimental survey [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2014, 36(7): 1442-1468.
- [2] ROSS D A, LIM J, LIN R S, *et al.*. Incremental learning for robust visual tracking [J]. *International Journal of Computer Vision*, 2008, 77(1-3): 125-141.
- [3] KWON J, LEE K M. Visual tracking decomposition [C]. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010: 1269-1276.
- [4] BABENKO B, YANG M H, BELONGINE S. Robust object tracking with online multiple instance learning [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, 33(8): 1619-1632.
- [5] HARE S, SAFFARI A, TORR P H S. Struck: Structured output tracking with kernels [C]. *IEEE International Conference on Computer Vision (ICCV)*, 2011: 263-270.
- [6] KALAL Z, MIKOLAJCZYK K, MATAS J. Tracking-learning-detection [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34(7): 1409-1422.
- [7] ZHANG K, ZHANG L, YANG M H. Real-time compressive tracking [C]. *European Conference on Computer Vision (ECCV)*, 2012: 864-877.
- [8] 陈东成,朱明,高文,等. 在线加权多实例学习实时目标跟踪[J]. *光学精密工程*, 2014, 22(6): 1661-1667.
- CHEN D CH, ZHU M, GAO W, *et al.*. Real-time object tracking via online weighted multiple instance learning [J]. *Opt. Precision Eng.*, 2014, 22(6): 1661-1667. (in Chinese)
- [9] 程帅,孙俊喜,曹永刚,等. 增量深度学习目标跟踪[J]. *光学精密工程*, 2015, 23(4): 1161-1170.
- CHEN SH, SUN J X, CAO Y G, *et al.*. Target tracking based on incremental deep learning [J]. *Opt. Precision Eng.*, 2015, 23(4): 1161-1170. (in Chinese)
- [10] 修春波,魏世安. 显著性直方图模型的 Camshift 跟踪方法[J]. *光学精密工程*, 2015, 23(6): 1750-1757.
- XIU CH B, WEI SH AN. Camshift tracking with saliency histogram [J]. *Opt. Precision Eng.*, 2015, 23(6): 1750-1757. (in Chinese)
- [11] 郭敬明,何昕,魏仲慧. 基于在线支持向量机的 Mean Shift 彩色图像跟踪[J]. *液晶与显示*, 2014, 29(1): 120-128.
- GUO J M, H X, W ZH H. New mean shift tracking for color image based on online support vector machine [J]. *Chinese Journal of Liquid Crystals and Displays*, 2014, 29(1): 120-128. (in Chinese)
- [12] 李静宇,王延杰. 基于子空间的目标跟踪算法研究[J]. *液晶与显示*, 2014, 29(4): 617-622.
- LI J Y, W Y J. Subspace based target tracking algorithm [J]. *Chinese Journal of Liquid Crystals and Displays*, 2014, 29(4): 617-622. (in Chinese)
- [13] BOLME D S, BEVERIDGE J R, DRAPER B A, *et al.*. Visual object tracking using adaptive correlation filters [C]. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010: 2544-2550.
- [14] HENRIQUES J F, CASEIRO R, MARTINS P, *et al.*. Exploiting the circulant structure of tracking-by-detection with kernels [C]. *European Conference on Computer Vision (ECCV)*, 2012: 702-715.
- [15] DANELLJAN M, KHAN F, FELSBERG M, *et al.*. Adaptive color attributes for real-time visual tracking [C]. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014: 1090-1097.
- [16] ZHANG K, ZHANG L, LIU Q, *et al.*. Fast visual tracking via dense spatio-temporal context learning [C]. *European Conference on Computer Vision (ECCV)*, 2014: 127-141.
- [17] HENRIQUES J F, CASEIRO R, MARTINS P, *et al.*. High-speed tracking with kernelized correlation filters [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(3): 583-596.
- [18] LI Y, ZHU J. A scale adaptive kernel correlation filter tracker with feature integration [C]. *European Conference on Computer Vision (ECCV)*,

- 2014; 254-265.
- [19] 张雷,王延杰,孙宏海,等. 采用核相关滤波器的自适应尺度目标跟踪[J]. 光学 精密工程, 2016, 24(2): 448-459.
- ZHANG L, WANG Y J, SUN H M, *et al.*. Adaptive scale object tracking with kernelized correlation filters [J]. *Opt. Precision Eng.*, 2016, 24(2): 448-459. (in Chinese)
- [20] 余礼杨,范春晓,明悦. 改进的核相关滤波器目标跟踪算法[J]. 计算机应用, 2015, 35(12): 3550-3554.
- YU L Y, FAN CH X, M Y. Improved target tracking algorithm based on kernelized correlation filter [J]. *Journal of Computer Application*, 2015, 35(12):3550-3554. (in Chinese)
- [21] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection [C]. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition(CVPR)*, 2005: 886-893.
- [22] WU Y, LIM J, YANG M H. Online object tracking: A benchmark [C]. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition(CVPR)*, 2013: 2411-2418.
- [23] SCHOLKOPF B, SMOLA A J. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond* [M]. MIT press, 2002.
- [24] RIFKIN R, YEO G, POGGIO T. Regularized least-squares classification [J]. *Nato Science Series Sub Series III Computer and Systems Sciences*, 2003, 190: 131-154.

#### 作者简介:



蔡玉柱(1990—),男,山东日照人,硕士研究生,2013年于天津职业技术师范大学获得学士学位,主要从事目标检测、目标跟踪方面的研究。E-mail: caiyuzhu001@sina.com

#### 导师简介:



杨德东(1977—),男,辽宁阜新人,副教授,硕士生导师,2000年、2003年于大连铁道学院分别获得学士、硕士学位,2007年于东北大学获得博士学位,主要从事智能感知与控制、目标检测与跟踪等方面的研究。E-mail: ydd12677@163.com

(版权所有 未经许可 不得转载)