

## 基于样本质量估计的空间正则化自适应相关滤波视觉跟踪

侯志强<sup>①②</sup> 王帅<sup>\*①②</sup> 廖秀峰<sup>③</sup> 余旺盛<sup>③</sup> 王姣尧<sup>③</sup> 陈传华<sup>③</sup>

<sup>①</sup>(西安邮电大学计算机学院 西安 710121)

<sup>②</sup>(西安邮电大学陕西省网络数据分析与智能处理重点实验室 西安 710121)

<sup>③</sup>(空军工程大学电讯工程学院 西安 710077)

**摘要:** 相关滤波(CF)方法应用于视觉跟踪领域中效果显著,但是由于边界效应的影响,导致跟踪效果受到限制,针对这一问题,该文提出一种基于样本质量估计的正则化自适应的相关滤波视觉跟踪算法。首先,该算法在滤波器的训练过程中加入空间惩罚项,构建目标与背景的颜色及灰度直方图模板并计算样本质量系数,使得空间正则项根据样本质量系数自适应变化,不同质量的样本受到不同程度的惩罚,减小了边界效应对跟踪的影响;其次,通过对样本质量系数的判定,合理优化跟踪结果及模型更新,提高了跟踪的可靠性和准确性。在OTB2013和OTB2015数据平台上的实验数据表明,与近几年主流的跟踪算法相比,该文算法的成功率均为最高,且与空间正则化相关滤波(SRDCE)算法相比分别提高了9.3%和9.9%。

**关键词:** 视觉跟踪; 相关滤波; 正则化自适应; 样本质量估计

中图分类号: TP391

文献标识码: A

文章编号: 1009-5896(2019)08-1983-09

DOI: 10.11999/JEIT180921

## Adaptive Regularized Correlation Filters for Visual Tracking Based on Sample Quality Estimation

HOU Zhiqiang<sup>①②</sup> WANG Shuai<sup>\*①②</sup> LIAO Xiufeng<sup>③</sup> YU Wangsheng<sup>③</sup>  
WANG Jiaoyao<sup>③</sup> CHEN Chuanhua<sup>③</sup>

<sup>①</sup>(School of Computer Science & Technology, Xi'an University of Posts & Telecommunications, Xi'an 710121, China)

<sup>②</sup>(Shaanxi Key Laboratory of Network Data Analysis and Intelligent Processing, Xi'an University of Posts and Telecommunications, Xi'an 710121, China)

<sup>③</sup>(Institute of Information and Navigation, Air Force Engineering University, Xi'an 710077, China)

**Abstract:** Correlation Filters (CF) are efficient in visual tracking, but their performance is badly affected by boundary effects. Focusing on this problem, the adaptive regularized correlation filters for visual tracking based on sample quality estimation are proposed. Firstly, the proposed algorithm adds spatial regularization matrix to the training process of the filters, and constructs color and gray histogram templates to compute the sample quality factor. Then, the regularization term adaptively changes with the sample quality coefficient, so that the samples of different quality are subject to different degrees of punishment. Then, by thresholding the sample quality coefficient, the tracking results and model update strategy are optimized. The experimental results on OTB2013 and OTB2015 indicate that, compared with the state-of-the-art tracking algorithm, the average success ratio of the proposed algorithm is the highest. The success ratio is raised by 9.3% and 9.9% contrasted with Spatially RegularizeD Correlation Filters(SRDCE) algorithm respectively on OTB2013 and OTB2015.

**Key words:** Visual tracking; Correlation Filters (CF); Adaptive regularization; Sample quality estimation

收稿日期: 2018-09-27; 改回日期: 2019-05-20; 网络出版: 2019-05-27

\*通信作者: 王帅 2289010261@qq.com

基金项目: 国家自然科学基金(61473309, 61703423)

Foundation Items: The National Natural Science Foundation of China (61473309, 61703423)

## 1 引言

视觉跟踪是计算机视觉领域及相关领域的一大研究热点<sup>[1,2]</sup>,在视频监控<sup>[3]</sup>、无人机制导等军民领域广泛应用。近年来,随着研究的深入以及新特征和新观点的提出,涌现出大量跟踪算法,视觉跟踪的性能显著提升,所面临的困难也逐渐得到解决。但是在实际跟踪中,目标受到复杂背景及自身变化等因素的影响,设计一个准确快速的跟踪系统仍然是一个困难的任务。

相关滤波(Correlation Filters, CF)的跟踪算法是视觉跟踪的研究重点,快速的跟踪速度使得其具有很大的改进空间。Bolme等人<sup>[4]</sup>将相关滤波引入视觉跟踪中,提出了误差最小平方和滤波(Minimum Output Sum of Squared Error filter, MOSSE)算法,只需要一个样本训练滤波器,并通过离散傅里叶变换计算所有候选样本与目标的相似度,其跟踪速度可以达到约700帧/s;Henriques等人<sup>[5]</sup>提出了核相关滤波(Kernelized Correlation Filters, KCF)算法,采用方向梯度直方图(HOG)特征<sup>[6]</sup>代替了灰度特征,并用循环采样的方法解决了采样的难题,跟踪准确性大大提升;Danelljan等人<sup>[7]</sup>提出颜色命名相关滤波(Color-Name, CN)算法,将降维后的颜色特征加入滤波器,进一步提高了针对彩色视频的跟踪能力;Zhang等人<sup>[8]</sup>提出了时空上下文相关滤波(Spatio-Temporal Context, STC)算法,利用深度时空上下文信息,提取目标周围的环境信息加入滤波器中进行训练,提高了算法针对遮挡的性能;Ma等人<sup>[9]</sup>将使用卷积网络层提取的目标特征与相关滤波跟踪框架融合,提出深层卷积特征跟踪算法(Hierarchical Convolutional Features, HCF)算法,提高了跟踪精度。Nam等人<sup>[10]</sup>使用大量具有标签值的视频序列训练网络得到普适性的目标模型,提出MDnet结构,解决训练数据不足的问题,Bertinetto等人<sup>[11]</sup>引入Siamese网络结构,训练得出完全端到端的跟踪模板,同时输入样本与候选样本,通过离线训练评估相似度。Danelljan等人<sup>[12]</sup>提出一个通用的框架来学习连续空间域中的卷积算子,提高了跟踪的准确度和鲁棒性。Ma等人<sup>[13]</sup>提出的长时相关滤波跟踪算法(Long-term Correlation filter Tracking, LCT),通过引入支持向量机(Support Vector Machine, SVM)分类器,当跟踪可能出现丢失时,提出对目标重新检测。Danelljan等人<sup>[14]</sup>提出使用尺度因子自适应针对目标尺度变化的相关滤波算法。文献<sup>[15,16]</sup>解决背景遮挡的相关滤波算法也相继被提出。

然而在相关滤波跟踪中,首先需要利用循环移

位获得大量样本训练滤波器,通过循环移位获取的样本中会有一部分样本出现边缘不平滑的问题即目标中心在图像的边缘,使用这样的样本训练会导致滤波器的辨别能力下降,出现模型漂移的可能大大增加。因此,为降低边界效应对跟踪的影响,一些空间正则化方法相继被提出, Danelljan等人<sup>[17]</sup>提出空间正则系数,空间正则化滤波器(Spatially Regularized Correlation Filters, SRDCF)算法根据空间位置对滤波器系数进行惩罚抑制,并通过高斯-赛德尔迭代算法(Gauss-Seidel algorithm)进行求解。Li等人<sup>[18]</sup>提出空间时间正则化相关滤波(Spatial-Temporal Regularized Correlation Filters, STRCF)算法,通过约束滤波器模板变化来改进SRDCF算法中多样本训练带来的计算负担,提高跟踪效果的同时加快了跟踪速度。然而STRCF算法中直接对所有训练滤波的样本采用完全相同的空间惩罚项,可能会错误抑制掉部分可信的目标响应;其次,空间正则化对背景信息有很大的抑制作用,而一定的背景的信息对跟踪是有利的,不能有效利用这些信息,对跟踪而言是一种损失。

针对边界效应对跟踪的影响,本文在STRCF算法的基础上提出一种基于样本质量估计的空间正则化自适应的相关滤波跟踪算法。本算法对多个视频进行分析:一是滤波器训练过程中,构建目标与背景的颜色及灰度统计直方图,定义质量系数对样本进行质量估计,通过改变空间正则化函数,寻找样本质量系数与正则化函数之间的联系,从而根据样本质量系数自适应改变空间正则化函数,使跟踪效果最优化;二是分析目标与背景统计信息的变化对跟踪效果的表达作用,通过对样本质量系数的阈值判定对正则化函数和模型更新进行优化,提高跟踪精度。

## 2 相关滤波模型

相关滤波方法的基本原理在于通过岭回归(ridge regression)对现有的样本以及标签进行训练,从而得到满足最小均方误差准则的滤波器参数,属于有监督的判别类跟踪方法。

### 2.1 相关滤波器的训练

相关滤波器利用岭回归对于任意输入都有简单的闭式最优解的特性,对现有的训练样本及标签 $(x_i, y_i), i = 1, 2, \dots, m$ 进行训练,寻找出对样本 $\mathbf{x} = (1 \ x_1 \ \dots \ x_n)^T$ 和标签值 $\mathbf{y} = (1 \ y_1 \ \dots \ y_m)^T$ 均方误差最小的滤波器参数 $\boldsymbol{\omega} = (\omega_1 \ \omega_2 \ \dots \ \omega_n)^T$ ,滤波器可以表示为 $f(\mathbf{x}) = \boldsymbol{\omega}^T \mathbf{x} + \xi$ ,  $\xi$ 表示预测值与真实值得偏差,通常服从正态分布 $\xi \propto N(0, \sigma^2)$ 。定义损失函数

$$J(\omega) = \min_{\omega} \sum_{i=1}^m L(y_i, f(x_i)) + \lambda \|\omega\|_2^2 \quad (1)$$

求极值得到滤波器参数

$$\omega = (\mathbf{x}^T \mathbf{x} + \lambda \mathbf{I})^{-1} \mathbf{x}^T \mathbf{y} \quad (2)$$

其中,  $\lambda$  是正则化参数, 避免分类器过拟合。根据 Representer 定理, 滤波器参数可表示为  $\omega = \sum_{i=1}^n \alpha_i \phi(x_i)$ 。其中, 核函数  $\phi(\mathbf{x})$  用于将训练样本映射到高维特征空间。定义核函数  $k(\mathbf{x}, \mathbf{x}') = \langle \phi(\mathbf{x}), \phi(\mathbf{x}') \rangle$ , 将滤波器参数求解问题转变为参数矩阵  $\alpha$  的求解过程:  $\alpha = (\mathbf{K}^{xx} + \lambda \mathbf{I})^{-1} \mathbf{y}$ , 其中  $\mathbf{K}^{xx}$  为核相关矩阵。根据循环矩阵性质:  $C(\mathbf{x}) \mathbf{x}' = F^{-1}(F^*(\mathbf{x}) \cdot F(\mathbf{x}'))$ , 可以将其变换到频域求解

$$\alpha = F^{-1} \left( \frac{1}{\hat{k}^{xx*} + \lambda} \cdot \hat{\mathbf{y}} \right) \quad (3)$$

其中,  $k^{xx} = k(\mathbf{X}, \mathbf{x}')$  为核相关矩阵  $\mathbf{K}^{xx}$  的第1行 ( $\mathbf{K}^{xx}$  也是循环矩阵),  $\hat{\cdot}$  表示变量的傅里叶变换式,  $\cdot$  表示元素的点积运算。

## 2.2 目标的检测

相关滤波器的检测过程是对候选样本进行分类的过程, 得到各个候选样本的响应值:  $f(\mathbf{Z}) = \langle \omega, \phi(\mathbf{Z}) \rangle$ , 其中  $\mathbf{Z}$  为样本  $\mathbf{z}$  循环得到的循环样本矩阵, 输入滤波器并在频域求解得到

$$f(\mathbf{Z}) = \left\langle \frac{\mathbf{y}}{k(\mathbf{X}, \mathbf{x}) + \lambda}, k^{xz} \right\rangle \quad (4)$$

核函数的选择, 可以选择高斯函数作为核函数

$$k^{xx} = \exp \left\{ - \left( \frac{\|\mathbf{x}\|^2 + \|\mathbf{x}'\|^2 - 2 \sum_{d=1}^D F^{-1}(\hat{\mathbf{x}}^* \cdot \hat{\mathbf{x}'})}{2\sigma^2} \right) \right\} \quad (5)$$

其中,  $D$  为样本的通道数, 也可以采用点积核函数

$$k^{xx} = F^{-1}(\hat{\mathbf{x}} \cdot \hat{\mathbf{x}'}) \quad (6)$$

高斯函数作为核函数可以得到较好的跟踪精度, 而速度较慢, 点积核函数精度略低一些, 而速度相对较快。

## 3 自适应正则化相关滤波跟踪

### 3.1 样本质量系数

由于受到目标运动、形变以及背景变化的影响, 当前帧样本与之前的样本相比, 会在一定程度上发生变化, 如何估计新样本的质量对于滤波器的训练以及后续的跟踪起到重要作用。本文针对样本质量的估计问题, 首先改进峰值旁瓣比(PSR)为

$$p = \frac{|R_{\max} - R_{\min}|^2}{\text{mean}(R_{ij} - \bar{R})^2}, \quad \forall i = 1, 2, \dots, m, j = 1, 2, \dots, n \quad (7)$$

其中,  $R$  为当前样本在滤波器中的响应,  $R_{\max}$ ,  $R_{\min}$  分别为响应的最大和最小值,  $R_{ij}$  为  $R$  中第  $i$  行第  $j$  列的元素,  $\bar{R}$  为响应均值。

系数  $p$  具有一定的评价样本质量的能力, 但是仅仅根据样本的响应进行样本质量估计是不全面的。因此, 本文为充分估计样本的质量, 考虑目标和背景的统计信息<sup>[19]</sup>, 提出了样本质量系数

$$Q = p \frac{e_t}{e_r} \quad (8)$$

其中,  $e_t$  与  $e_r$  分别为样本与样本模板、样本与背景之间的差异系数, 通过计算颜色(灰度)统计直方图的交叉熵得到

$$e_t = \sum_{i=1}^N h_{t-1_i} \log \frac{1}{h_{t_i}} \quad (9)$$

$$e_r = \sum_{i=1}^N H_{t_i} \log \frac{1}{h_{t_i}} \quad (10)$$

其中,  $N$  为直方图中 bin 的个数,  $h_{t-1}$  表示  $t-1$  帧中目标的统计直方图模板,  $h_t$  为当前目标的统计直方图,  $H_r$  为当前背景的统计直方图。

样本质量系数  $Q$  综合样本响应与样本及背景统计信息对样本进行质量估计。经过实验分析, 由图1可以看出: (1) 当样本发生形变时, 其响应会受到影响, 系数  $p$  减小, 此时通过样本与模板以及背景的统计信息的差异估计, 可以在一定程度上进行修正; (2) 而当目标受到背景影响时, 系数  $p$  的表达能力并不稳定, 而样本与背景信息之间的差异系数变小, 样本与模板间差异系数增大, 样本质量系数减小, 表明样本质量的降低; (3) 目标较长时间保持相似状态时, 说明该样本可信度较高, 质量系数随之升高; (4) 此外, 系数  $p$  对分辨率变化不敏感, 而质量系数  $Q$  对分辨率变化敏感, 当分辨率降低, 质量系数减小。

### 3.2 空间正则化自适应相关滤波

空间正则项的加入在于减小边界效应对滤波器训练的影响, 然而对所有样本采用相同惩罚值不免会错误抑制有效样本, 另外对背景信息的利用率也会降低。自适应调整空间正则项对滤波器训练有一定的益处。

本文通过3.1节中所提的样本质量系数  $Q$  对目标质量进行估计, 设定阈值  $\tau$  从而定义自适应空间正则项。当样本质量  $Q > \tau$  时, 采用自适应的反高斯函数定义

$$\omega = \varphi \cdot \exp \left( -0.5 \times \left( \frac{r^2 + c^2}{\delta_t^2} \right) \right) / Q \quad (11)$$

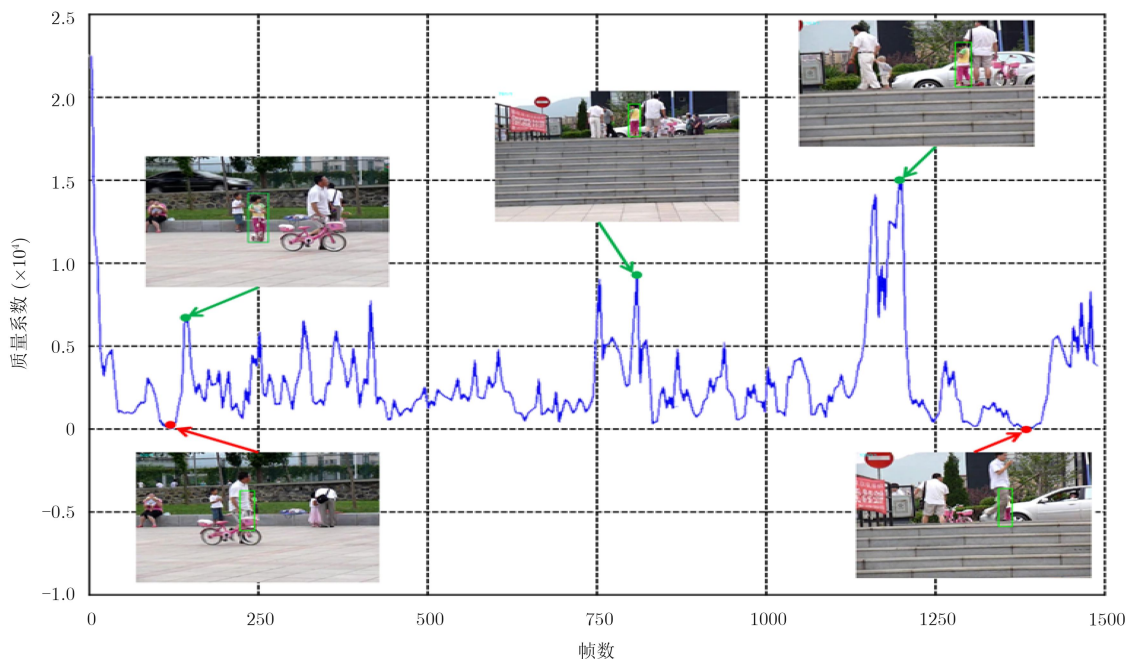


图1 质量系数曲线图

其中,  $\varphi$  为预设系数,  $\delta_t$  为标准差,  $r$  和  $c$  表示位置。

当样本质量  $Q < \tau$  时, 说明样本质量过低, 为避免背景信息的干扰, 直接采用赋值的空间正则项: 即在样本中, 目标所在位置为中心, 目标尺度为大小的区域赋最小值, 其余背景区域均赋最大值。

在STRCF算法的基础上, 加入样本质量的估计, 实现空间正则化自适应跟踪

$$\arg \min_{\omega} \left\| \sum_{d=1}^D \mathbf{x}_t^d \times \omega^d - \mathbf{y} \right\|^2 + \sum_{d=1}^D \|\mathbf{w} \cdot \omega^d\|^2 + \mu \|\omega - \omega_{t-1}\|^2 \quad (12)$$

其中,  $\omega_{t-1}$  为  $t-1$  帧中训练得到的滤波器参数,

$\mu$  为正则化系数,  $\sum_{d=1}^D \|\mathbf{w} \cdot \omega^d\|^2$  表示空间正则化,

$\|\omega - \omega_{t-1}\|^2$  表示时间正则化。

引入拉格朗日算子并利用交替方向乘法(Alternating Direction Method of Multipliers, ADMM)优化求解<sup>[16]</sup>, 可将其转换为若干子问题

$$\left. \begin{aligned} \omega^{(i+1)} &= \arg \min_{\omega} \left\| \sum_{d=1}^D \mathbf{x}_t^d \times \omega^d - \mathbf{y} \right\|^2 \\ &\quad + \gamma \|\omega - \mathbf{g} + \mathbf{h}\|^2 + \mu \|\omega - \omega_{t-1}\|^2 \\ \mathbf{g}^{(i+1)} &= \arg \min_{\mathbf{g}} \sum_{d=1}^D \|\mathbf{w} \cdot \mathbf{g}^d\|^2 \\ &\quad + \gamma \|\omega - \mathbf{g} + \mathbf{h}\|^2 \\ \mathbf{h}^{(i+1)} &= \mathbf{h}^{(i)} + \omega^{(i+1)} - \mathbf{g}^{(i+1)} \end{aligned} \right\} \quad (13)$$

其中,  $\mathbf{g}$  为引入的辅助变量且  $\mathbf{g} = \omega$ ,  $\mathbf{h} = \frac{1}{\gamma} \mathbf{s}$ ,  $\mathbf{s}$  为拉格朗日乘子,  $\gamma$  步长系数。

求解得出各子问题的解

$$\nu_j(\hat{\omega}) = \frac{1}{\mu + \gamma} \left( \mathbf{I} - \frac{\nu_j(\hat{\mathbf{x}}_t) \nu_j(\hat{\mathbf{x}}_t)^T}{\mu + \gamma + \nu_j(\hat{\mathbf{x}}_t)^T \nu_j(\hat{\mathbf{x}}_t)} \right) \mathbf{q} \quad (14)$$

其中,  $\mathbf{q} = \nu_j(\hat{\mathbf{x}}_t) \hat{\mathbf{y}}_j + \gamma \nu_j(\hat{\mathbf{g}}) - \gamma \nu_j(\hat{\mathbf{h}}) + \mu \nu_j(\hat{\omega}_{t-1})$ ,  $\nu_j(\hat{\omega})$  表示  $\hat{\omega}$  中的第  $j$  个元素。

$$\mathbf{g} = (\mathbf{W}^T \mathbf{W} + \gamma \mathbf{I})^{-1} (\gamma \omega + \gamma \mathbf{h}) \quad (15)$$

其中, 对角矩阵  $\mathbf{W}$  由  $D$  个对角矩阵  $\text{Diag}(\mathbf{w})$  构成。

步长参数  $\gamma$  的更新表示为

$$\gamma^{(i+1)} = \min(\gamma^{\max}, \rho \gamma^{(i)}) \quad (16)$$

其中,  $\gamma^{\max}$  为  $\gamma$  最大值,  $\rho$  为更新尺度。

### 3.3 模型更新策略

模型的更新主要体现在滤波器参数以及样本模板的更新, 从而减小跟踪过程中受到光照、目标形变、遮挡等因素影响, 出现跟踪漂移的现象。样本质量对模型更新有很重要的意义, 质量高的样本应着重利用。根据本文所提样本质量系数, 设定更新阈值  $\psi$ , 当样本质量  $Q > \psi$  时, 则通过式(17)进行滤波器模型更新以及目标统计模型的更新, 目标统计模型的更新

$$\mathbf{h}_t = (1 - \alpha) \mathbf{h}_{t-1} + \alpha \mathbf{h} \quad (17)$$

其中,  $\mathbf{h}_t$  表示当前帧更新后的直方图模板,  $\mathbf{h}_{t-1}$  表示更新前的直方图模板,  $\mathbf{h}$  为当前帧提取的目标直方图,  $\alpha$  为更新率。滤波器模型利用式(14)求解并更新。

当样本质量  $Q < \psi$  时，减小更新率，目标统计模型的更新

$$h_t = \left(1 - \frac{\alpha}{2}\right) h_{t-1} + \frac{\alpha}{2} h \quad (18)$$

滤波器模型更新中加入首帧信息

$$\omega_t = (1 - \alpha)\omega_{t-1} + \alpha\omega_1 \quad (19)$$

其中， $\omega_t$ ,  $\omega_{t-1}$ ,  $\omega_1$  分别表示更新后的滤波器参数、更新前的滤波器参数以及首帧信息训练得到的滤波器参数。

### 3.4 算法流程

综上对本文算法的描述，可将本文算法主要跟踪流程归纳为表1，算法的整体流程则如图2所示。

表 1 自适应正则化的相关滤波视觉跟踪算法

<b>输入：</b> 图像序列 $I_1, I_2, \dots, I_n$ ，目标初始位置 $p_0 = (x_0, y_0)$ ，目标初始尺度 $s_0 = (w_0, h_0)$ 。
<b>输出：</b> 每帧图像的跟踪结果，即目标位置 $p_t = (x_t, y_t)$ ，目标尺度估计 $s_t = (w_t, h_t)$
对于 $t = 1, 2, \dots, n$ ， <b>do</b> ：
(1) 目标定位及尺度估计
(a) 利用前一帧目标位置 $p_{t-1}$ 以及尺度 $s_{t-1}$ 确定第 $t$ 帧 ROI 区域；
(b) 提取多尺度样本 $I_s = \{I_{s_1}, I_{s_2}, \dots, I_{s_S}\}$ ；
(c) 根据响应图确定第 $t$ 帧中目标的中心位置 $p_t$ 以及尺度 $s_t$ ；
(2) 样本质量估计及正则化自适应
(a) 根据目标中心位置及尺度提取目标及背景统计直方图；
(b) 利用式(8)计算样本质量系数 $Q$ ；之后，利用样本质量系数计算空间正则化项；
(3) 模型更新
(a) 利用式(19)更新跟踪滤波器模型 $\omega_t$ ；
(b) 利用式(17)、式(18)更新统计信息模型 $h_t$ ；
<b>结束</b>

## 4 仿真实验

为充分验证本文算法的有效性，本文在Windows操作系统下，采用MATLAB 2014a进行编程实现。将本文算法在Intel Xeon 2.4 GHz处理器上进行测试，并采用GPU(TITANX)进行加速，算法参数设置如下：多尺度采样尺度个数  $S=7$ ，采样尺度因子  $a=1.01$ ，感兴趣的背景区域为目标大小的2倍，反高斯正则化系数  $\varphi=10^9$ ，方差为  $\delta=0.1$ ，时间正则化系数  $\mu=16$ ，步长尺度因子  $\rho=1.2$ ，统计模板更新率  $\alpha=0.025$ ，本文算法所提空间正则化阈值  $\tau$  决定着样本质量的好坏的判定，因此选取合适的自适应空间正则化阈值，可提高滤波器的判别力，不合适的阈值选取反而影响滤波器的判别力。表2为阈值  $\tau=2500, 2750, 3000, 3250, 3500, 3750$  时，本文算法在OTB2015中的结果，结果表明当阈值  $\tau$  选择为3000时，效果最好。模型更新阈值  $\psi=1000$ 。以上参数均为大量实验得出的经验值，并在所有的实验中保持不变。

为了充分说明本文算法在应对边界效应以及背景干扰等方面的优势，本文在OTB2013<sup>[20]</sup>数据集上与9种当前主流算法进行对比。以及在OTB2015<sup>[21]</sup>数据集上与8种当前主流的跟踪算法进行对比，结果如图3所示。选取的对比算法有STRCF, HCF<sup>[22]</sup>, STCT<sup>[23]</sup>, CNN-SVM<sup>[24]</sup>, DSST<sup>[11]</sup>, KCF, TGPR<sup>[25]</sup>, C-COT<sup>[9]</sup>, SMC<sup>[13]</sup>，跟踪结果由其作者提供。

### 4.1 定性分析

(1) 目标遮挡：目标遮挡使得跟踪算法不能获得足够的目标信息而导致跟踪偏差或丢失。以“Gir12”和“Soccer”为例，目标在跟踪过程中受到背景遮挡，在“Gir12”第100~120帧中均受到较为严重的遮挡，在此期间得到的样本质量较低，

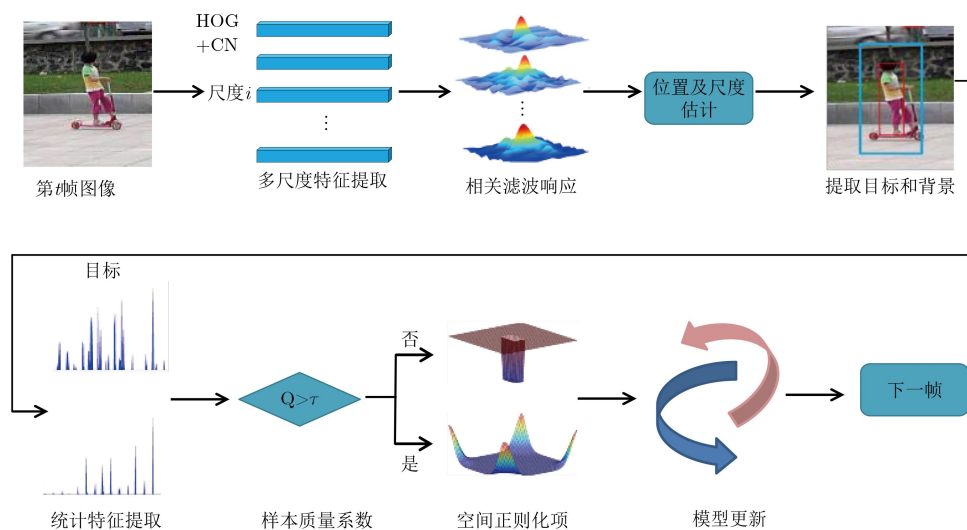


图 2 本文算法流程图

表 2 阈值 $\tau$ 的选取与OTB2015实验结果的对比分析

阈值 $\tau$	2500	2750	3000	3250	3500	3750
OTB2015跟踪成功率	0.820	0.779	0.871	0.855	0.817	0.795

本文算法根据质量系数在训练和更新上的改进使得当目标重新出现能够准确将其捕获，而其它算法无法重新寻找到目标，导致效果降低。

(2) 目标快速运动：目标外观发生变化会导致当前帧信息与模板匹配程度降低，增加跟踪难度，特别是对于相关滤波而言，边界效应会导致边缘的样本割裂，快速运动的目标难以得到较好的响应。例如“Bolt2”和“KiteSurf”，目标的快速运动和形变增加了跟踪难度，本文算法考虑背景信息对于跟踪的利害，对其进行自适应的抑制，可以较稳定地对目标进行跟踪。

(3) 目标形变：目标外观发生变化会导致当前帧信息与模板匹配程度降低，增加跟踪难度。以“Sylvester”和“Basketball”为例，目标在跟踪过程中不断发生形态变化或者空间里的旋转，导致模板难以一直保持较高的匹配，部分算法会出现跟踪漂移的现象，而本文算法对样本质量的估计中加入统计信息，统计信息对形态变化更加鲁棒，在正确估计出当前样本的质量的情况下，目标定位及尺度估计更加准确，取得了较好的训练和跟踪效果。

(4) 尺度变化：尺度变化是目标跟踪中比较常

见的问题，由于目标尺度的变化导致固定大小的跟踪框无法获取到完整的目标信息从而影响跟踪效果甚至丢失目标。本文算法对模板更新策略上的阈值判断抑制了坏样本的影响，在目标重新出现后仍然能够很快找到目标并估计其尺度。

4.2 定量分析

为充分分析算法的有效性，本文分别从单个视频和数据集整体两个方面对算法进行定量分析。

(1) 单个视频性能分析

采用平均中心位置误差和成功率对所选取的8组视频进行对比分析。表3为本文算法(AWDCF)与9种对比跟踪算法对所选的8组视频的跟踪成功率和平均中心误差对比，括号前的数字为平均中心误差(像素)，括号内数字的为覆盖率为0.5时的成功率(%)。其中标红的数字为9种算法中的最优，蓝色字体数字为次优。可以看出对于大部分视频而言，本文算法均保持较高的成功率和较低的平均中心误差，跟踪性能良好。

(2) 算法综合性能分析

分别在文献[20,21]中给出的两个数据集上进行试验和对比测试，采用跟踪成功率来评价算法综合性能。成功率为跟踪结果中覆盖率大于指定阈值的帧数占整个视频帧数的百分比，而精度为算法跟踪结果的平均中心误差在指定阈值以内的帧数占整个视频帧数的百分比。



图 3 8种算法的部分跟踪结果对比

图4为在OTB2015数据集和OTB2013数据集上得到的跟踪算法的成功率曲线，可以看出，本文算法的成功率位于前列，在OTB2013数据集上和OTB2015数据集上分别为0.693和0.666。相比于STRCF跟踪算法，本文算法的成功率均有一定的提升；另外，本文算法的跟踪速度(18.8 帧/s)，比采

用深度学习的HCF算法(10.2 帧/s)和SMCF(8.5 帧/s)更加快速，也比同样采用相关滤波的SRDCF算法(5.8 帧/s)跟踪速度更快；而由于增加了样本质量系数的计算，比STRCF算法的跟踪速度(24.3 帧/s)略有下降。

表4分别为10种算法在不同属性的视频中的成

表 3 8组测试序列的中心误差(像素)和成功率(%)

算法	CNN-SVM	STRCF	TGPR	HCF	KCF	STCT	DSST	C-COT	SMCF	
Girl2	7.6(98.0)	11.3(89.0)	30.9(87.0)	110.0(56.0)	118.8(8.0)	264.6(7.0)	319.1(8.0)	46.4(54.0)	8.4(96.0)	7.9(97.0)
Soccer	17.5(81.0)	260(24.0)	19.6(62.0)	60.7(14.0)	13.5(53.0)	15.6(46.0)	46.9(18.0)	14.3(43.0)	12.1(83.0)	14.5(84.0)
Bolt2	6.4(90.0)	151.4(48.0)	7.8(71.0)	304.0(1.0)	8.3(88.0)	329.8(1.0)	6.3(95.0)	115.5(1.0)	7.0(92.0)	6.8(90.0)
KiteSurf	2.3(99.0)	25.2(51.0)	66.7(45.0)	61.7(38.0)	59.8(45.0)	40.6(31.0)	7.8(70.0)	56.7(43.0)	2.1(99.0)	2.3(99.0)
Sylvester	5.5(96.0)	5.0(98.0)	5.5(96.0)	5.7(91.0)	12.9(83.0)	13.3(81.0)	14.8(82.0)	14.8(70.0)	4.5(99.0)	7.5(99.0)
Basketball	3.8(99.0)	21.4(48.0)	14.1(11.0)	9.4(90.0)	3.7(100.0)	8.1(90.0)	3.9(98.0)	111.6(14.0)	5.0(97.0)	4.1(98.0)
Dog1	3.0(100.0)	7.2(58.0)	3.6(100.0)	5.9(69.0)	4.4(67.0)	4.1(64.0)	4.7(97.0)	4.6(66.0)	4.0(98.0)	4.8(96.0)
CarScale	7.4(77.0)	19.8(53.0)	8.7(72.0)	21.4(46.0)	29.3(73.0)	16.1(55.0)	15.2(77.0)	18.8(51.0)	5.3(87.0)	8.7(77.0)
平均	5.8(94.0)	51.7(58.0)	16.2(74.2)	70.6(62.0)	26.6(62.0)	71.3(50.8)	42.5(57.0)	38.9(47.0)	5.4(93.9)	6.1(93.5)

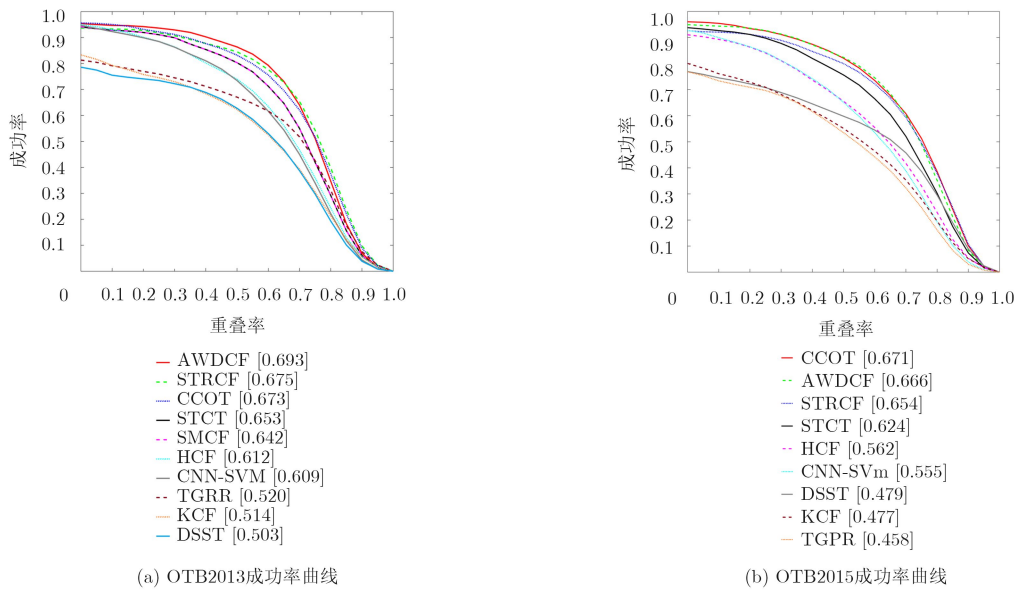


图 4 OTB2013和OTB2015数据集上的精度和成功率曲线

表 4 不同属性下算法跟踪成功率对比结果

	IV (40)	OPR (64)	SV (66)	OCC (50)	DEF (44)	MB (31)	FM (41)	IPR (31)	OV (14)	BC (33)	LR (10)
本文算法	0.659	0.644	0.640	0.641	0.624	0.672	0.646	0.622	0.600	0.655	0.570
CNN-SVM	0.532	0.546	0.492	0.513	0.547	0.568	0.530	0.545	0.488	0.543	0.419
STRCF	0.646	0.628	0.637	0.618	0.607	0.666	0.634	0.604	0.585	0.639	0.561
TGPR	0.449	0.454	0.400	0.429	0.412	0.409	0.398	0.461	0.373	0.426	0.378
HCF	0.535	0.532	0.487	0.523	0.530	0.573	0.555	0.557	0.474	0.575	0.424
KCF	0.469	0.449	0.399	0.438	0.436	0.456	0.452	0.464	0.393	0.489	0.306
STCT	0.636	0.584	0.596	0.592	0.603	0.625	0.616	0.570	0.530	0.625	0.527
DSST	0.476	0.448	0.414	0.426	0.412	0.465	0.442	0.484	0.374	0.463	0.311
C-COT	0.641	0.637	0.654	0.639	0.637	0.688	0.610	0.635	0.613	0.666	0.583
SMCF	0.672	0.653	0.632	0.653	0.612	0.665	0.632	0.610	0.608	0.663	0.579

功率对比, 其中, IV, OPR, SV, OCC, DEF, MB, FM, IPR, OV, BC, LR分别为光照变化属性、平面向外旋转属性、尺度变化属性、遮挡属性、形变属性、运动模糊属性、快速运动属性、平面内旋转属性、视野外属性、背景复杂属性、低分辨率属性。本文算法对各种属性的视频的跟踪成功率在10种算法中表现良好, 特别是针对目标快速移动和被遮挡的视频跟踪效果明显优于其他算法。

## 5 结束语

本文提出一种基于样本质量估计的空间正则化自适应的相关滤波算法。本算法针对相关滤波中的边界效应问题, 提出利用样本响应以及目标和背景统计信息计算样本质量系数, 得到自适应变化的空间正则项, 实现了对不同质量样本自适应地采用不同的正则项来抑制背景信息, 在一定程度上解决了边界效应的问题。同时, 对于目标和背景统计信息的利用使得算法更加准确, 对模板更新和跟踪也具有一定的帮助。试验结果表明, 本文算法在复杂环境下具有良好的稳定性, 在一定程度上解决了边界效应对跟踪的影响。

## 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. doi: [10.1109/TPAMI.2013.230](https://doi.org/10.1109/TPAMI.2013.230).
- [2] WANG Naiyan, SHI Jianping, YEUNG D Y, *et al.* Understanding and diagnosing visual tracking systems[C]. Proceedings of IEEE International Conference on Computer Vision, Santiago, Chile, 2015: 3101–3109.
- [3] 黄立勤, 朱飘. 车载视频下改进的核相关滤波跟踪算法[J]. 电子与信息学报, 2018, 40(8): 1887–1894. doi: [10.11999/JEIT171109](https://doi.org/10.11999/JEIT171109).  
HUANG Liqin and ZHU Piao. Improved kernel correlation filtering tracking for vehicle video[J]. *Journal of Electronics & Information Technology*, 2018, 40(8): 1887–1894. doi: [10.11999/JEIT171109](https://doi.org/10.11999/JEIT171109).
- [4] BOLME D S, BEVERIDGE J R, DRAPER B A, *et al.* Visual object tracking using adaptive correlation filters[C]. Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, USA, 2010: 2544–2550.
- [5] HENRIQUES J F, CASEIRO R, MARTINS P, *et al.* High-speed tracking with kernelized correlation filters[J]. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 2015, 37(3): 583–596. doi: [10.1109/TPAMI.2014.2345390](https://doi.org/10.1109/TPAMI.2014.2345390).
- [6] FELZENSZWALB P F, GIRSHICK R B, MCALLESTER D, *et al.* Object detection with discriminatively trained part-based models[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, 32(9): 1627–1645. doi: [10.1109/TPAMI.2009.167](https://doi.org/10.1109/TPAMI.2009.167).
- [7] DANELLJAN M, KHAN F S, FELSBERG M, *et al.* Adaptive color attributes for real-time visual tracking[C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Columbus, USA, 2014: 1090–1097.
- [8] ZHANG Kaihua, ZHANG Lei, LIU Qingshan, *et al.* Fast visual tracking via dense spatio-temporal context learning[C]. The 13th European Conference on Computer Vision, Zurich, Switzerland, 2014: 127–141.
- [9] MA Chao, HUANG Jiabin, YANG Xiaokang, *et al.* Hierarchical convolutional features for visual tracking[C]. IEEE International Conference on Computer Vision, Santiago, Chile, 2015: 3074–2082.
- [10] NAM H and HAN B. Learning multi-domain convolutional neural networks for visual tracking[C]. IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, USA, 2016: 4293–4302.
- [11] BERTINETTO L, VALMADRE J, HENRIQUES J F, *et al.* Fully-convolutional siamese networks for object tracking[C]. Computer Vision – ECCV 2016 Workshops, Amsterdam, the Netherlands, 2016: 850–865.
- [12] DANELLJAN M, ROBINSON A, KHAN F S, *et al.* Beyond correlation filters: Learning continuous convolution operators for visual tracking[C]. The 14th European Conference on Computer Vision, Amsterdam, The Netherlands, 2016: 472–488.
- [13] MA Chao, YANG Xiaokang, ZHANG Chongyang, *et al.* Long-term correlation tracking[C]. IEEE Conference on Computer Vision and Pattern Recognition, Boston, USA, 2015: 5388–5396.
- [14] DANELLJAN M, HAGER G, KHAN F S, *et al.* Discriminative scale space tracking[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(8): 1561–1575. doi: [10.1109/TPAMI.2016.2609928](https://doi.org/10.1109/TPAMI.2016.2609928).
- [15] LI Feng, YAO Yingjie, LI Peihua, *et al.* Integrating boundary and center correlation filters for visual tracking with aspect ratio variation[C]. IEEE International Conference on Computer Vision Workshops, Venice, Italy, 2017: 2001–2009.
- [16] WANG Xin, HOU Zhiqiang, YU Wangsheng, *et al.* Online scale adaptive visual tracking based on multilayer convolutional features[J]. *IEEE Transactions on Cybernetics*, 2019, 49(1): 146–158. doi: [10.1109/TCYB.2017.2768570](https://doi.org/10.1109/TCYB.2017.2768570).
- [17] DANELLJAN M, HÄGER G, KHAN F S, *et al.* Learning

- spatially regularized correlation filters for visual tracking[C]. IEEE International Conference on Computer Vision, Santiago, Chile, 2015: 4310–4318.
- [18] LI Feng, TIAN Cheng, ZUO Wangmeng, *et al.* Learning spatial-temporal regularized correlation filters for visual tracking[C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA, 2018: 4904–4913.
- [19] 毕笃彦, 库涛, 查宇飞, 等. 基于颜色属性直方图的尺度目标跟踪算法研究[J]. 电子与信息学报, 2016, 38(5): 1099–1106. doi: [10.11999/JEIT150921](https://doi.org/10.11999/JEIT150921).  
BI Duyan, KU Tao, ZHA Yufei, *et al.* Scale-adaptive Object tracking based on color names histogram[J]. *Journal of Electronics & Information Technology*, 2016, 38(5): 1099–1106. doi: [10.11999/JEIT150921](https://doi.org/10.11999/JEIT150921).
- [20] WU Yi, LIM J, and YANG M H. Online object tracking: A benchmark[C]. IEEE Conference on Computer Vision and Pattern Recognition, Portland, USA, 2013: 2411–2418.
- [21] 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).
- [22] MA Chao, HUANG Jiabin, YANG Xiaokang, *et al.* Hierarchical convolutional features for visual tracking[C]. Proceedings of IEEE International Conference on Computer Vision, Santiago, Chile, 2015: 3074–3082.
- [23] WANG Lijun, OUYANG Wanli, WANG Xiaogang, *et al.* STCT: sequentially training convolutional networks for visual tracking[C]. IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, USA, 2016: 1373–1381.
- [24] 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 Machine Learning, Lille, France, 2015: 597–606.
- [25] GAO Jin, LING Haibin, HU Weiming, *et al.* Transfer learning based visual tracking with Gaussian processes regression[C]. The 13th European Conference on Computer Vision, Zurich, Switzerland, 2014: 188–203.
- 侯志强: 男, 1973年生, 教授, 研究方向为计算机视觉和模式识别.
- 王 帅: 男, 1995年生, 硕士生, 研究方向为计算机视觉和机器学习.
- 廖秀峰: 男, 1993年生, 硕士生, 研究方向为计算机视觉和机器学习.
- 余旺盛: 男, 1985年生, 讲师, 研究方向为计算机视觉和图像处理.
- 王皎尧: 女, 1995年生, 硕士生, 研究方向为计算机视觉和机器学习.
- 陈传华: 男, 1994年生, 硕士生, 研究方向为计算机视觉和机器学习.