

## 基于自适应多模型联合的目标跟踪算法

王任华<sup>1</sup>, 沈剑宇<sup>2</sup>, 蒋敏<sup>2</sup>

(1. 中国人民公安大学 信息技术与网络安全学院, 北京 100038; 2. 江南大学 物联网工程学院, 江苏 无锡 214122)

**摘要:** 基于核相关滤波器的跟踪算法对于目标的空间结构具有较强的依赖性, 无法有效应对遮挡、形变等干扰因素, 且单一的特征模型在复杂的跟踪场景下无法准确表述目标信息。为此, 提出一种基于自适应多模型联合的算法。通过自适应权重将相关滤波模型与颜色直方图模型进行联合, 并将稀疏表示的思想引入相关滤波模型的训练过程中, 以增强算法的鲁棒性。OTB 视频序列数据集上的实验结果表明, 该算法可有效缓解跟踪过程中的遮挡、形变等因素的干扰, 与 Staple 算法、KCF 算法相比, 目标跟踪的精度显著提升。

**关键词:** 目标跟踪; 自适应权重; 核相关滤波器; 联合模型;  $L_1$  范数

开放科学(资源服务)标志码(OSID):

中文引用格式: 王任华, 沈剑宇, 蒋敏. 基于自适应多模型联合的目标跟踪算法[J]. 计算机工程, 2019, 45(8): 266-274.

英文引用格式: WANG Renhua, SHEN Jianyu, JIANG Min. Target tracking algorithm based on adaptive multi-model joint[J]. Computer Engineering, 2019, 45(8): 266-274.

## Target Tracking Algorithm Based on Adaptive Multi-Model Joint

WANG Renhua<sup>1</sup>, SHEN Jianyu<sup>2</sup>, JIANG Min<sup>2</sup>

(1. College of Information Technology and Network Security, People's Public Security University of China, Beijing 100038, China;  
2. School of Internet of Things Engineering, Jiangnan University, Wuxi, Jiangsu 214122, China)

**[Abstract]** The tracking algorithm based on Kernelized Correlation Filter(KCF) has strong dependence on the spatial structure of the target, cannot effectively deal with the occlusion, deformation and other interference factors, and the single feature model is unable to accurately represent the target information in complex tracking occasions. Therefore, an adaptive multi-model tracking algorithm is proposed. The correlation filter model and the color histogram model are combined by adaptive weight, and the idea of sparse representation is introduced into the training process of the correlation filter model to further enhance the robustness of the algorithm. Experimental results on the OTB dataset show that the algorithm can effectively mitigate the interferences from various factors such as occlusion and deformation in the tracking process, and compared with Staple algorithm, KCF algorithm and others, the algorithm achieves significant improvement in target tracking accuracy.

**[Key words]** target tracking; adaptive weight; Kernelized Correlation Filter(KCF); joint model;  $L_1$  norm

DOI: 10.19678/j.issn.1000-3428.0052593

### 0 概述

目标跟踪作为机器视觉领域最具挑战性的研究内容之一, 受到国内外研究者的关注和重视, 其在视频监控、无人驾驶、智能诊断、人机交互等领域有不同程度的应用<sup>[1]</sup>。但在实际复杂的环境中, 跟踪目标时会出现多种干扰因素, 包括目标遮挡、目标形

变、外部光照变化、背景杂乱等<sup>[2]</sup>。这些干扰因素会导致非目标像素的累积, 产生漂移现象, 最终跟踪失败<sup>[3]</sup>。近年来, 相关滤波器因其计算速度和定位性能在目标跟踪领域取得许多成功的应用<sup>[4]</sup>。文献[5]将相关滤波器应用到目标跟踪领域, 提出基于最小输出平方和误差(Minimum Output Sum of Squared Error, MOSSE)的相关滤波器。文献[6]在此基础上

**基金项目:** 国家自然科学基金(61362030, 61201429); 中国博士后科学基金(2015M581720, 2016M600360); 科技援疆专项计划(2017E0279); 公安部技术研究计划(2018JSYJA01)。

**作者简介:** 王任华(1972—), 女, 副教授、硕士, 主研方向为视频侦查、模式识别; 沈剑宇, 硕士研究生; 蒋敏, 教授、博士。

**收稿日期:** 2018-09-07 **修回日期:** 2018-10-24 **E-mail:** renhuawang@163.com

提出基于循环结构的核检测跟踪(Circulant Structure of tracking-by-detection with Kernel, CSK)算法,利用循环移位进行密集采样并引入岭回归。然而,CSK 算法仅采用灰度特征,缺乏对目标信息的完整表述。文献[7]将核方法引入 CSK 中,并与梯度方向直方图(Histogram of Oriented Gradient, HOG)<sup>[8]</sup>特征结合,提出核相关滤波器(Kernelized Correlation Filter, KCF),其在速度和精度上都取得较好的效果。KCF 算法对空间结构的依赖性导致其对形变、遮挡等干扰敏感<sup>[9]</sup>。当目标发生形变时,目标以及背景的整体颜色分布并不会随之改变,因此,基于颜色直方图的跟踪算法可以很好地应对形变等干扰因素<sup>[10]</sup>的影响,但这类方法易受背景中相似颜色物体的干扰,对运动模糊和光照变化等因素较为敏感。文献[11]提出颜色感知算法(Distractor-Aware Tracker, DAT),通过自适应阈值来抑制与目标颜色相似的背景的影响,但对于背景混杂的情况,其判别性与基于相关滤波器的跟踪算法相比有一定的欠缺。文献[12]在此基础上提出 Staple 算法,通过分配固定的权重将 DAT 算法与基于核相关滤波器的算法进行联合学习,取得较好的效果,但是固定权重的联合在复杂情况下无法体现其优势。

在上述研究的基础上,本文提出基于自适应多模型联合的跟踪算法,将 KCF 跟踪模型和基于颜色直方图的跟踪模型进行联合。在目标位置评估时,采用自适应权重联合的策略,充分利用 2 种模型的互补特性,实现鲁棒的目标跟踪。由于在跟踪过程中,多种干扰因素会对目标外观产生影响,本文在 KCF 模型训练的过程中加入  $L1$  范数的正则化<sup>[13]</sup>,利用稀疏跟踪的部分思想,使滤波器能更好地拟合目标期望值,以提高算法的精确度。

## 1 基于核相关滤波器的目标跟踪

KCF 算法在候选区域内通过对基样本循环移位采样获取大量训练样本,对于得到的训练样本与标签  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ , 利用岭回归对滤波器  $w$  进行训练,使得样本  $x_i$  与高斯型的期望值  $y_i$  之间的平方误差最小并获得闭合解。KCF 算法的损失函数如下:

$$\min_w \sum_i (w^T x_i - y_i)^2 + \lambda \|w\|_2^2 \quad (1)$$

其中,  $\lambda$  为惩罚项。上式的矩阵形式可以表示为:

$$\min_w \|Xw - y\|_2^2 + \lambda \|w\|_2^2 \quad (2)$$

其中,  $X = [x_1, x_2, \dots, x_n]^T$  为训练样本矩阵。对  $w$  求导,并令导数为 0,得如下解:

$$w = (X^T X + \lambda I)^{-1} X^T y \quad (3)$$

为了将上式转换到傅里叶域内,实现快速计算,先将上式转化为复数域表达:

$$w = (X^H X + \lambda I)^{-1} X^H y \quad (4)$$

其中,  $X^H = (X^*)^T$  为  $X$  的复共轭转置矩阵,  $X^*$  表示  $X$  的复共轭。一般通过求解高维线性方程得到式(4)的解, KCF 算法可以根据循环矩阵的特性,利用离散傅里叶变换(Discrete Fourier Transform, DFT)来避免求解高维线性方程,节省计算时间。此时,  $X$  可表示为如下形式:

$$X = F \text{diag}(\hat{x}) F^H \quad (5)$$

其中,  $\hat{x}$  为 DFT 变换后的结果,  $F$  为离散傅里叶矩阵,是常量。由此,式(4)可以表示为:

$$w = F \text{diag}\left(\frac{\hat{x}^*}{\hat{x}^* \odot \hat{x} + \lambda}\right) F^H y \quad (6)$$

其中,  $\odot$  表示按像素点相乘。利用循环矩阵的卷积性质将上式简化为:

$$\hat{w} = F(w) = \frac{\hat{x}^* \odot \hat{y}}{\hat{x}^* \odot \hat{x} + \lambda} \quad (7)$$

其中,  $F$  表示 DFT。由此时域中的解  $w$  可以通过傅里叶逆变换得到,避免复杂矩阵的求逆等运算。对一幅  $m \times m$  的图像实行卷积操作,其时间复杂度为  $O(m^4)$ ,通过 DFT 变换将卷积操作转化到频域中,则相同情形下所需的时间复杂度仅为  $O(m^2 \lg m)$ ,计算速率得到大幅提升。

考虑到实际应用中大量的非线性可分数据,文献[7]引入核方法,利用高低维空间之间的映射关系,将参数转换到对偶空间,则  $w$  可由映射后的样本  $\varphi(x)$  的线性组合<sup>[14]</sup>表示:  $w = \sum_i \alpha_i \varphi(x_i)$ , 由此得到岭回归的解<sup>[15]</sup>:

$$\alpha = (K + \lambda I)^{-1} y \quad (8)$$

其中,  $\alpha$  是由  $\alpha_i$  构成的向量,  $K$  为核矩阵,由核函数  $k(x, x') = \varphi^T(x) \varphi(x')$  组成,  $I$  为单位矩阵。无论是线性核、高斯核还是多项式核,大部分的核函数都满足循环结构。因而,可以利用循环矩阵在傅里叶域的特殊性质来求解  $\alpha$ :

$$\hat{\alpha} = \frac{\hat{y}}{\hat{k}^{xx} + \lambda} \quad (9)$$

其中,  $\hat{k}^{xx}$  为核矩阵  $\mathbf{K}$  的首行元素。最后, 对于任意测试样本  $\mathbf{Z}$  的滤波器响应值可以表示为:

$$f(\mathbf{Z}) = \mathbf{w}^T \mathbf{z} = \sum_i^{n-1} \alpha_i k(z, x_i) \quad (10)$$

## 2 自适应多特征联合模型的目标跟踪算法

基于颜色直方图的算法模型利用目标与背景的颜色统计信息, 可以很好地应对目标遮挡与形变等情况, 但无法在光照变化和运动模糊等条件下取得较好成绩。基于核相关滤波器的跟踪模型由于刚性模板的固有特性, 其对目标形变非常敏感, 为弥补这一问题, 本文通过自适应权重联合 2 个模型, 提高算法在复杂情况下的鲁棒性。

### 2.1 基于 L1 范数的核相关滤波器跟踪模型

在跟踪目标的过程中, 由于目标外观受到遮挡、光照变化、物体形变等因素的影响, 在取样的过程中有较大误差, 产生离群值影响。简单的平方损失函数无法很好地拟合这种情形, 容易引起过拟合, 导致跟踪失败。为此, 本文在文献[16]的 KCF 损失函数的基础上, 参考稀疏表示, 引入 L1 范数以优化原始的目标函数。通过 L1 损失允许跟踪过程中出现稀疏离群值的特性, 使得训练得到的滤波器可以很好地拟合高斯型目标回归值, 可在很大程度上忽略跟踪过程中由目标外观变化带来的较大误差值, 从而缓解过拟合现象, 获得更加鲁棒的目标函数。具体形式如下:

$$\min_{\mathbf{w}, \mathbf{e}} \|\varphi(\mathbf{X})\mathbf{w} - \mathbf{y} + \mathbf{e}\|_2^2 + \lambda \|\mathbf{w}\|_2^2 + \beta \|\mathbf{e}\|_1 \quad (11)$$

其中,  $\mathbf{X}$  为由循环采样构成的样本矩阵,  $\varphi(\cdot)$  为映射函数,  $\mathbf{y}$  是期望值,  $\lambda$  与  $\beta$  均为正则化参数。在式(11)中的  $\mathbf{e} = \mathbf{y} - \varphi(\mathbf{X})\mathbf{w}$  为残差值,  $\|\mathbf{e}\|_1$  项用来消除离群值的影响, 正则化项  $\|\mathbf{w}\|_2^2$  用来保证回归的稳定性, 提高模型的可靠性。

上述的目标函数无法直接求解, 但可将其划分成 2 个子问题, 通过迭代计算获得式(11)的近似解:

$$\min_{\mathbf{w}} \|\varphi(\mathbf{X})\mathbf{w} - \mathbf{y} + \mathbf{e}\|_2^2 + \lambda \|\mathbf{w}\|_2^2 \quad (12)$$

$$\min_{\mathbf{e}} \|\varphi(\mathbf{X})\mathbf{w} - \mathbf{y} + \mathbf{e}\|_2^2 + \beta \|\mathbf{e}\|_1 \quad (13)$$

在求解式(12)时,  $\mathbf{e}$  取某个固定值  $\mathbf{e}'$ , 得到:

$$\min_{\mathbf{w}} \|\varphi(\mathbf{X})\mathbf{w} - \mathbf{y} + \mathbf{e}'\|_2^2 + \lambda \|\mathbf{w}\|_2^2 \quad (14)$$

式(14)与 KCF 算法中的原始损失函数类似, 可根据 KCF 算法的求解方法获得其在对偶空间的解, 如下:

$$\hat{\alpha} = \frac{\hat{\mathbf{y}} - \hat{\mathbf{e}}}{\hat{\mathbf{k}}^{xx} + \lambda} \quad (15)$$

最后通过逆傅里叶变换得到当前最优解  $\alpha'$ 。

类似地, 在求解式(13)时, 将当前的最优解  $\alpha'$  固定, 然后求解  $\mathbf{e}$  的最优值, 优化问题可以改写为:

$$\min_{\mathbf{e}} \|\mathbf{K}\alpha' - \mathbf{y} + \mathbf{e}\|_2^2 + \beta \|\mathbf{e}\|_1 \quad (16)$$

式(16)的求解作为一个凸优化问题, 可以通过收缩阈值算子<sup>[17]</sup>进行求解:

$$\mathbf{e} = S_{\beta}(\mathbf{F}^{-1}(\hat{\mathbf{y}} - \hat{\alpha}\mathbf{e}\hat{\mathbf{k}}^{xx})) \quad (17)$$

其中,  $S$  为收缩阈值算子, 计算如下:

$$S_{\beta}(x) = \text{sign}(x) \max(|x| - \beta, 0) \quad (18)$$

在此帧的候选区域内, 即以上一帧的目标位置为中心的采样框内, 通过循环采样获取样本  $z$ , 对其检测获取响应值, 如下:

$$S_w = f(z) = \mathbf{F}^{-1}(\hat{\mathbf{k}}^{xz} \mathbf{e} \hat{\alpha}) \quad (19)$$

### 2.2 基于颜色直方图的跟踪模型

为缓解形变等干扰的负面效应, 本文在候选区域  $\Omega$  上引入基于贝叶斯分类器  $h$  的颜色直方图模型, 其响应值可以表示为积分图像的平均值, 如下:

$$S_h = \mathbf{w}^T \left( \frac{1}{|\Omega|} \sum_{\gamma \in \Omega} I[\gamma] \right) \quad (20)$$

其中,  $\mathbf{w}$  表示该模型的权重矩阵,  $I$  表示区域  $\Omega$  内颜色直方图特征。对于一个区域  $\Omega'$ , 该模型的损失函数表示为:

$$\min_{\mathbf{w}} \frac{1}{|\Omega'|} \sum_{\gamma \in \Omega'} (\mathbf{w}^T I[\gamma] - y)^2 \quad (21)$$

其中,  $y$  为回归目标值。式(21)可以化简为在目标区域  $O$  和目标以外区域  $B$  分别提取样本  $(v, 1)$  和  $(v', 0)$ , 得到如下形式:

$$\min_{\mathbf{w}} \frac{1}{|O|} \sum_{\gamma \in O} (\mathbf{w}^T I[\gamma] - 1)^2 + \frac{1}{|B|} \sum_{\gamma \in B} (\mathbf{w}^T I[\gamma])^2 \quad (22)$$

上述线性回归方程的解为:

$$\mathbf{w} = \frac{\rho(O)}{\rho(O) + \rho(B)} \quad (23)$$

其中,  $\rho(H) = N(H)/(H)$ ,  $N(H)$  是直方图特征总数,  $N(H) = |\{\gamma \in H: I[\gamma] \neq 0\}|$ 。

### 2.3 模型联合位置估计

基于颜色直方图的算法模型不依赖于图像特征的空间排列, 虽能提高在目标形变和遮挡情况下的

算法鲁棒性,但当目标与背景颜色相近时,该算法不具备较强的判别能力,难以将目标从周围环境中区分出来。基于核相关滤波器的跟踪算法可较好地目标从复杂背景中分离出来,但该算法对图像特征的空间分布很敏感,当目标外观发生较大变化时,其性能急剧降低。针对 2 种模型各自的特点,本文采用自适应联合权重的方法联合这 2 个模型,提高算法在复杂情况下的鲁棒性。最后的目标位置可以由联合响应图得到:

$$S = \psi_h S_h + \psi_w S_w \quad (24)$$

其中,  $\psi_h$  和  $\psi_w$  分别是核模型与颜色模型的权重,且  $\psi_w = 1 - \psi_h$ 。固定比例的方法虽然可以省去一些不必要的运算,但在不同环境下的普适性有所欠缺。

根据 KCF 算法的实现原理以及在实验过程中每帧的最大响应值变化可知,当目标外观受到剧烈变化时,该帧的响应值会急剧变小。如图 1 所示,在 Bolt 视频序列的第 254 帧目标开始受到遮挡物影响,此时,该帧由 KCF 算法得到的最大响应值为 0.269,到下一帧时,目标已经被完全遮挡,得到的最大响应值下降为 0.194。

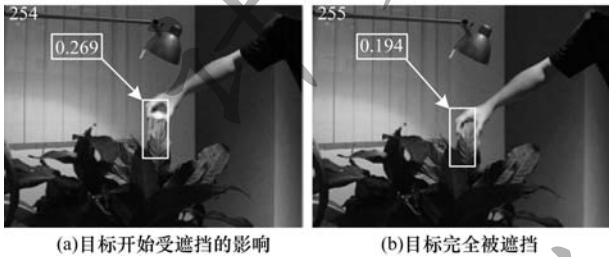


图 1 KCF 算法的响应情况

基于对上述实验规律的总结与分析,为保证联合模型的有效性,本文提出自适应权重的联合标准,根据 2 个模型各自的响应值来确定响应的比例,如下:

$$\psi_h = \frac{\max(S_h)}{\max(S_h) + \max(S_w)} \quad (25)$$

其中,  $\max(S_h)$ 、 $\max(S_w)$  分别表示模型在当前帧得到的响应图的最大值。由前文可知,当目标受遮挡、形变等因素影响时,  $\max(S_w)$  会突然变小,但颜色模型由于对形变不敏感,其响应图的  $\max(S_h)$  基本维持稳定状态。若当前帧出现遮挡情况,改进 KCF 模型的响应图的可信度下降,  $\max(S_w)$  减小,模型最终

的响应值更依赖于颜色模型,因此,通过式(25)可以得到较大的  $\psi_h$ ,而  $\psi_w$  的值较小。最终,通过式(24)可以得到联合响应图,其最大值位置即为预测的当前帧目标位置。

## 2.4 尺度估计

为了适应跟踪过程中目标尺度的变化,本文基于文献[18]的思想,在跟踪算法的框架中使用独立的核相关滤波器作为尺度滤波器,对目标尺度的变化进行处理。当由联合响应图得到此帧的目标位置后,以该位置为中心,围绕中心取不同尺度大小的图像块提取特征,分别与尺度滤波器进行相关操作以获得响应图,响应值最大的响应图所对应的样本尺度即为当前帧的目标尺度。

设  $A \times B$  表示前一帧的目标尺度,首先围绕当前帧的目标位置提取大小为  $a^n A \times a^n B$  的尺度样本  $j^n$ ,其中,  $n \in \left\{ \left\lfloor -\frac{S-1}{2} \right\rfloor, \left\lfloor -\frac{S-2}{2} \right\rfloor, \dots, \left\lfloor \frac{S-1}{2} \right\rfloor \right\}$ ,  $S$  表示尺度等级,  $a$  为尺度因子。通过式(26)计算每个尺度样本的响应值,具体计算过程如下:

$$S_{\text{scale}} = F^{-1} \left\{ \frac{\sum_{k=1}^d j_n^k \overline{M^k}}{N + \gamma} \right\} \quad (26)$$

其中,  $d$  是特征维度,  $\delta$  为惩罚项。由此得到的最大响应值所对应的目标尺寸即为当前帧的目标尺度。式(26)中的  $M$  和  $N$  可通过式(27)和式(28)得到:

$$M_t = (1 - \theta) M_{t-1} + \theta \overline{G}_t F_t \quad (27)$$

$$N_t = (1 - \theta) N_{t-1} + \theta \sum_{k=1}^d \overline{F^k} F^k \quad (28)$$

其中,  $t$  表示帧,  $\theta$  表示学习率。

## 2.5 算法框架

本文跟踪算法的框架如图 2 所示。当跟踪第  $t$  帧的目标时,首先在候选窗内循环移位得到训练样本,提取 HOG 特征,并通过 Cosine 窗平滑边缘噪声。然后,与优化后的基于  $L1$  范数正则化的核相关滤波器进行相关操作得到响应图,同时,由颜色直方图模型得到同一候选窗内的直方图响应。通过自适应权重对 2 个模型进行有效融合,将融合后的响应图最大值的位置作为当前帧的目标位置,并提取 HOG 特征。与独立的尺度滤波器进行相关操作后,得到该帧的目标尺度信息,从而确定第  $t$  帧的目标状态。最后,对整个算法模型的滤波器进行更新。

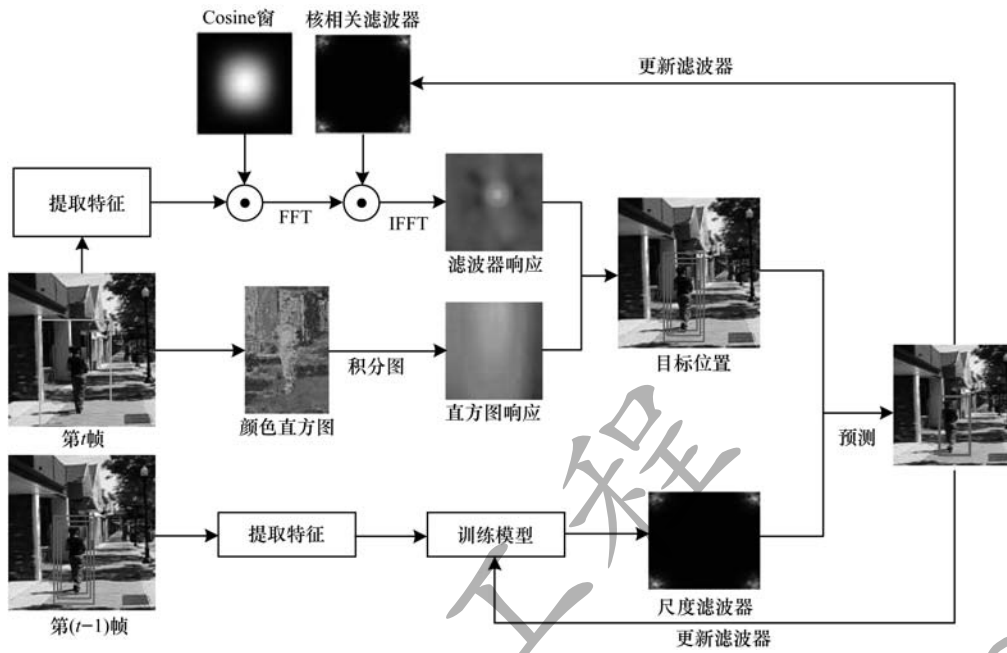


图2 本文跟踪算法框架

### 3 实验结果与分析

本文的实验环境为 Intel Core i5 CPU@3.20 GHz, Window7 系统, 内存 8 GB, 所用的软件平台为 Matlab R2014a。

#### 3.1 实验参数

与其他核相关滤波跟踪模型类似, 本文使用 HOG 特征进行实验。在位置估计时, 式(11)中的  $\lambda$  与  $\beta$  的取值均为  $1 \times 10^4$ , 求解式(11)的迭代次数设置为 15 次。此外, 联合权重的初始值采用  $\psi_h$  和  $\psi_w$  分别取 0.3 和 0.7。在尺度估计时, 式(26)中的  $\delta$  取值为  $1 \times 10^4$ , 等级  $S$  取 33, 尺度因子  $a$  为 1.02,  $\theta$  设置为 0.025。

#### 3.2 评估方法

本文实验采用文献[19]提出的 OTB 视频数据集, 并沿用该视频数据集的评估指标: 准确率和成功率。其中, 准确率由预测的目标中心位置与手工标定的准确位置之间的平均欧氏距离, 即中心像素误差 (Center Location Error, CLE) 来确定。当某一帧中心位置误差小于给定阈值时, 认为该算法在这一帧成功, 并以成功跟踪的帧数占总帧数的百分比与阈值的关系绘制精确度图。成功率以重叠率  $S$  为评价指标, 其具体计算过程如下:  $S = \frac{|r_i \cap r_a|}{|r_i \cup r_a|}$ , 其中,  $r_a$  和  $r_i$  分别表示得到的目标跟踪框和手工标定的准确跟踪框。若  $S$  大于给定阈

值, 则认为跟踪成功。以重叠率  $S$  与阈值的关系绘制覆盖成功率图。

将本文算法与 OTB 数据集上的 25 个经典算法<sup>[19]</sup>, 即 IVT、LIAPG、SemiT、CT、CPF、Frag、VTS、DFT、VR-V、BSBT、TLD、ORIA、ASLA、MTT、LSK、CXT、LOT、SMS、OAB、MIL、SCM、VTD、KMS、Struck、CSK, 以及 KCF<sup>[7]</sup>、DSST<sup>[18]</sup>、MTSA<sup>[20]</sup>、Staple<sup>[12]</sup> 算法进行对比实验。其中包括生成式跟踪算法和判别式跟踪算法, 同时将拥有类似思想的 KCF 算法和 Staple 算法列入对比实验组。上述所有算法的代码从 OTB 数据库和相应作者的主页上获得, 并选取最佳参数进行实验, 且所有结果均在同一实验平台下完成。实验使用的视频序列涵盖了数据库标定的 11 种干扰因素, 分别为形变、遮挡、光照变化、运动模糊、尺度变化、平面内旋转、平面外旋转、快速移动、背景杂乱、超出视野范围和低分辨率。

#### 3.3 结果分析

考虑到视频初始化位置对跟踪算法的影响, 本文采用 OTB 数据集中的时间鲁棒性评价 (Temporal Robustness Evaluation, TRE) 指标进行评估, 即利用视频序列中的任意帧作为实验的第 1 帧进行跟踪, 实验的准确率和成功率对比结果分别如图 3、图 4 所示, 图例中方括号内的值为 AUC (Area Under Curve) 值。为便于观察, 图中仅显示排名前 10 的算法结果。从图 3、图 4 可以看出, 除了在快速移动这一干扰因素

下,本文算法(Ours\_co 算法)效果略逊于 Staple 外,在其余干扰因素下本文算法都表现出良好的鲁棒性,在所有对比算法中位列第一。尤其是在应对形变和光照干扰方面,本文算法与排名第 2 的 Staple 算法相比,准确率和成功率都提高了 5% 左右。在遮挡干扰的处理上,本文算法也表现出较好的鲁棒性。同时,本文

跟踪算法在 OTB 数据库上的平均速度为 58.3 frame/s,完全可以满足实际应用环境下的实时性要求。为进一步体现本文算法的应用价值,选取精度前 10 名的算法做实时性分析,结果如表 1 所示。从表 1 可以看出,本文算法的平均速度为 58.3 frame/s,可以满足实际应用环境下的实时性要求。

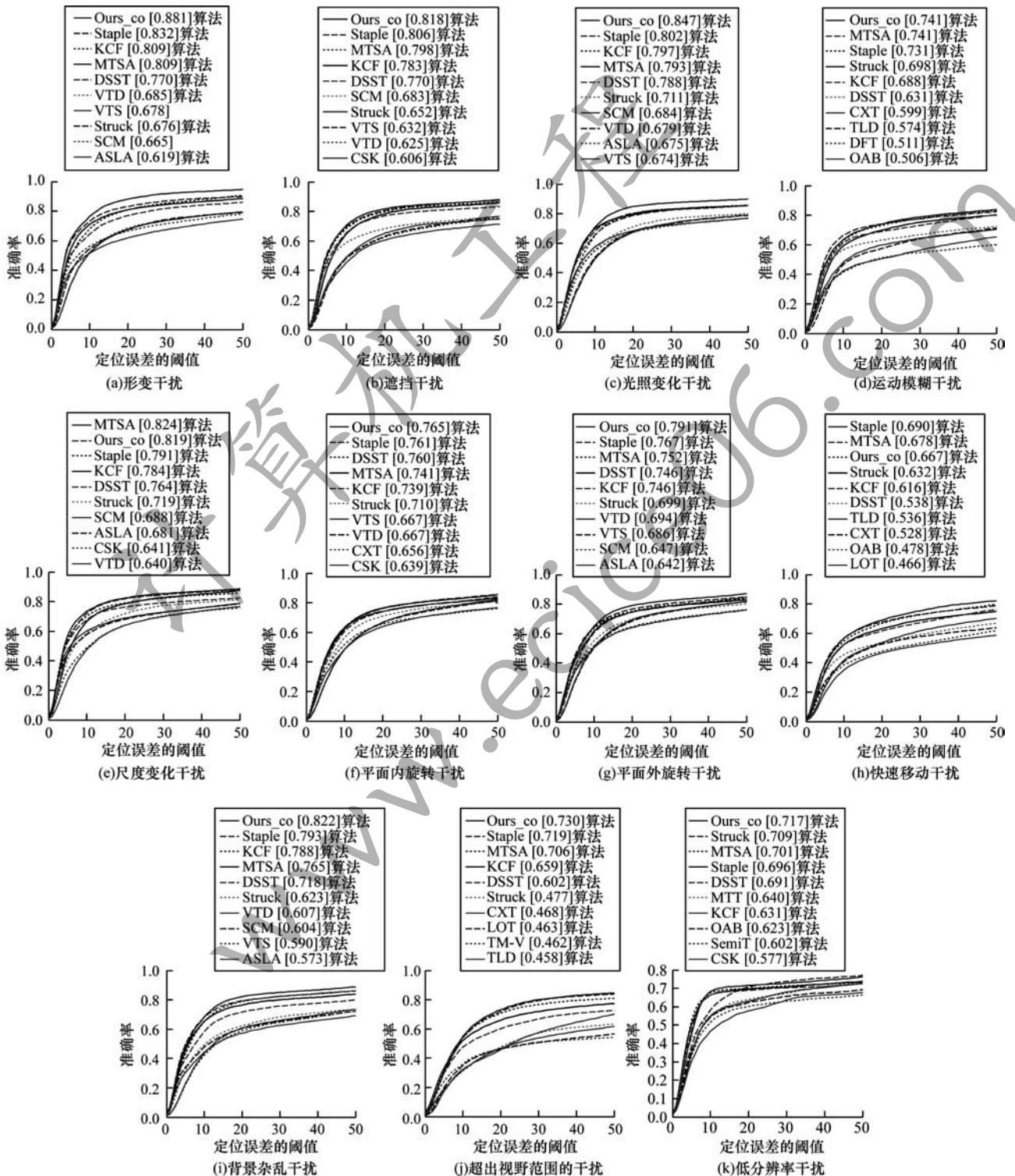


图 3 主流跟踪算法在不同干扰下的准确率对比

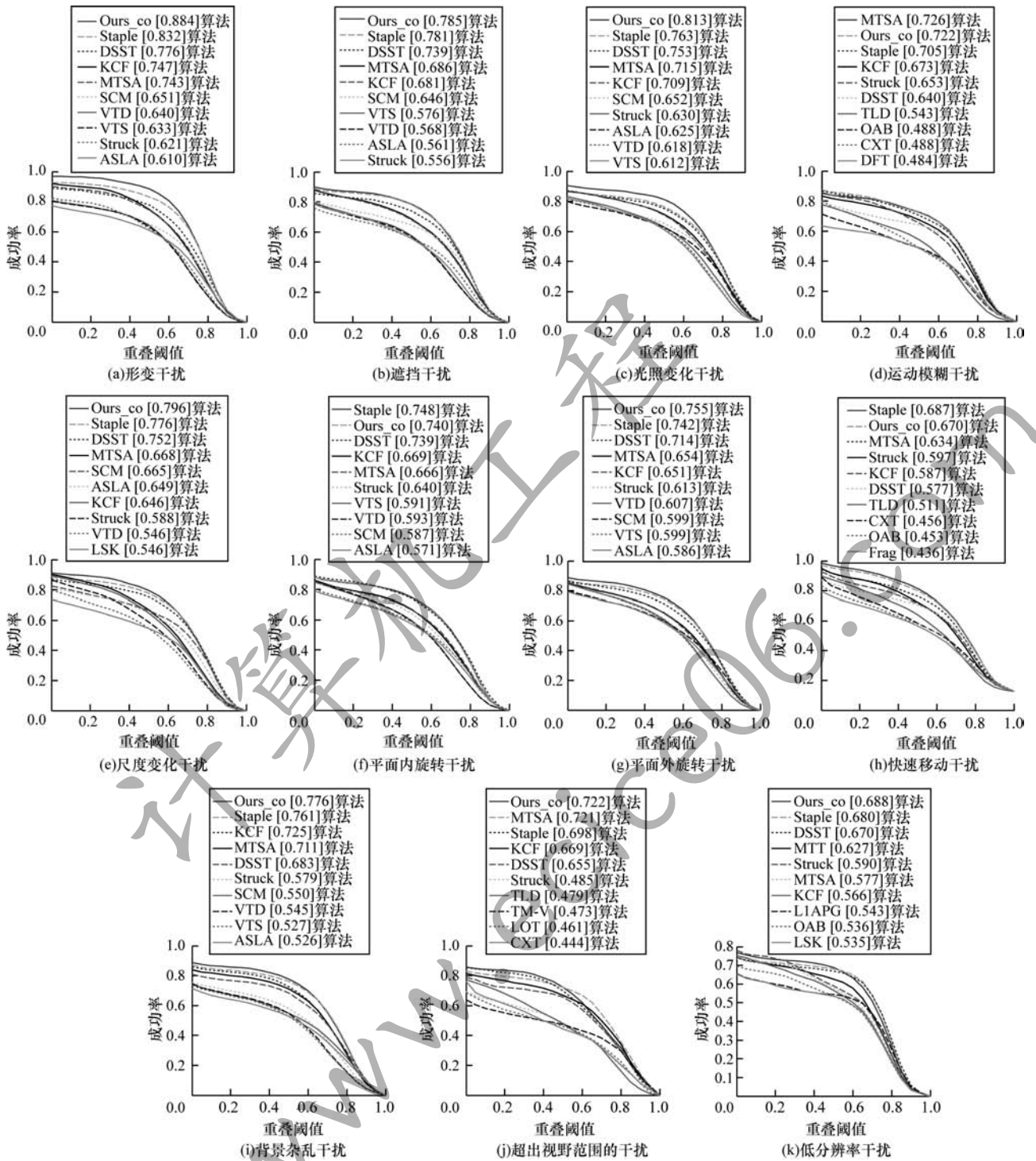


图4 主流跟踪算法在不同干扰下的成功率对比

表1 精度前10名算法的速度对比 (frame · s<sup>-1</sup>)

算法	速度	算法	速度
Ours_co	58.3	Staple	44.9
KCF	189.1	DSST	25.9
MTSA	20.5	Struck	10.1
VTS	2.3	VTD	2.3
ASLA	4.2	SCM	0.2

为更直观地展现算法在不同视频序列上、不同干扰因素下的跟踪效果,本文从 OTB 数据库中选取 6 个具有代表性的视频序列,其包含的干扰因素如表 2 所示。图 5 给出本文算法与 KCF 算法、Staple 算法在这 6 个视频序列上的对比跟踪效果。其中,实线、虚线和点划线框出的部分分别对应本文算法、KCF 算法和 Staple 算法的结果。

表 2 6 个视频序列的干扰因素

视频序列	干扰因素
Basketball	光照变化、遮挡、形变、旋转、背景杂乱
CarScale	尺度变化、遮挡、快速运动、旋转
FaceOcc2	光照变化、遮挡、旋转
Couple	尺度变化、遮挡、形变、快速运动、旋转、背景杂乱
Trellis	光照变换、尺度变化、背景杂乱、旋转
Skating1	光照变化、尺度变化、遮挡、形变、旋转、背景杂乱

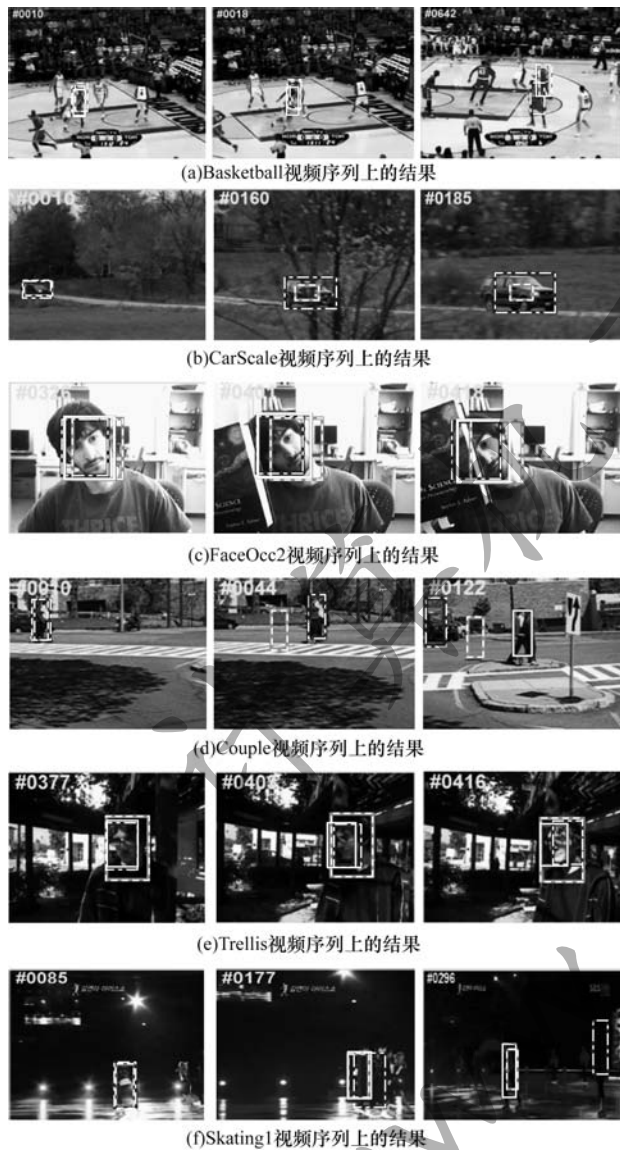


图 5 3 种算法在 6 个视频序列上的跟踪结果对比

从图 5 可以看出,在 Basketball 序列中,第 10 帧时 3 种算法都能实现较准确的跟踪,随着遮挡逐渐严重,KCF 算法和 Staple 算法开始出现目标漂移。当发生形变、光照变化时,Staple 算法定位的目标位置完全偏离真实目标位置,这是因为 Staple 算法在实现模型联合时,只是简单地分配固定比例的模型权重,当相关滤波模型在产生较大偏差却又被分配较高权重时,跟踪失败。本文算法根据每一帧的响应值适时地调整权重比例,当 2 个模型中任何一个

出现较大偏差时,通过式(25)可自适应地分配权重,缓解干扰因素对最终跟踪结果的影响。在 CarScale 序列中,本文算法对尺度变化展现出良好的应对能力。在 FaceOcc2 序列中,当目标发生旋转和遮挡时,KCF 算法和 Staple 算法都发生不同程度的中心位置偏移。相比较而言,本文算法表现出较好的准确性。在 Couple 序列中,背景杂乱、遮挡、形变、快速运动等多种干扰因素的影响较严重。在第 44 帧,受到背景杂乱的干扰,KCF 算法在训练过程中引入过多的错误训练样本,滤波器区分背景和目标的能力下降,最终跟踪失败。随着跟踪的推进,错误像素慢慢累积,产生较大误差,而 Staple 算法无法有效利用互补模型对其进行矫正,最终跟踪失败,而本文算法实现较为稳定、准确的跟踪。在 Trellis 序列中,Staple 算法逐渐发生较大的中心偏差,KCF 算法未对目标尺度进行处理,而本文算法始终能够成功定位到目标(脸部)。在 Skating1 序列中,第 177 帧的目标受到相似颜色的物体遮挡,颜色模型的响应图出现偏差,但由于未能实时调整权重比例,Staple 算法开始发生漂移,最终跟踪失败。KCF 算法虽然能大致定位到目标,但无法应对目标的尺度变化,也出现了目标中心偏移的现象。

#### 4 结束语

本文提出一种自适应的多模型联合算法。将基于 L1 范数的核相关滤波器跟踪模型与颜色直方图模型相结合,利用 2 个模型各自的良好特性,以解决形变等干扰因素导致的跟踪失败问题,优化其在复杂环境下的鲁棒性。通过自适应联合权重来调整 2 个模型的权重比例,最终确定目标位置。实验结果表明,相较于其他生成式跟踪算法和判别式跟踪算法,该算法在复杂环境下具有更好的鲁棒性。下一步将简化跟踪算法的目标函数,并尝试寻找更加高效的模型融合策略,以提高跟踪结果的精度,减少算法的运行时间。

#### 参考文献

[1] 张晶,王旭,范洪博. 自适应学习的时空上下文目标跟踪算法[J]. 计算机工程,2018,44(6):294-299.  
 [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. Washington D. C., USA:IEEE Press,2015:3101-3109.  
 [3] 孔军,成静,蒋敏,等. 组合范数正则化稀疏编码和自适应加权残差的鲁棒跟踪[J]. 计算机辅助设计与图形学学报,2018,30(4):634-641.

- [ 4 ] 魏全禄,老松杨,白亮. 基于相关滤波器的视觉目标跟踪综述[J]. 计算机科学,2016,43(11):1-5.
- [ 5 ] BOLME D S, BEVERIDGE J R, DRAPER B A, et al. Visual object tracking using adaptive correlation filters[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press,2010:2544-2550.
- [ 6 ] HENRIQUES J F, CASEIRO R, MARTINS P, et al. Exploiting the circulant structure of tracking-by-detection with kernels[C]//Proceedings of European Conference on Computer Vision. Berlin, Germany: Springer, 2012: 702-715.
- [ 7 ] HENRIQUES J F, RUI C, MARTINS P, et al. High-speed tracking with kernelized correlation filters [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2014,37(3):583-596.
- [ 8 ] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection [ C ]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2005:886-893.
- [ 9 ] YAO Rui,XIA Shixiong,ZHANG Zhen,et al. Real-time correlation filter tracking by efficient dense belief propagation with structure preserving [ J ]. IEEE Transactions on Multimedia,2017,19(4):772-784.
- [ 10 ] 李冠彬,吴贺丰. 基于颜色纹理直方图的带权分块均值漂移目标跟踪算法[J]. 计算机辅助设计与图形学学报,2011,23(12):2059-2066.
- [ 11 ] POSSEGGER H, MAUTHNER T, BISCHOF H. In defense of color-based model-free tracking [ C ]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press,2015:2113-2120.
- [ 12 ] BERTINETTO L, VALMADRE J, GOLODETZ S, et al. Staple:complementary learners for real-time tracking[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2016:1401-1409.
- [ 13 ] BAO Chenglong, WU Yi, LING Haibin, et al. Real time robust L1 tracker using accelerated proximal gradient approach [ C ]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA:IEEE Press,2012:1830-1837.
- [ 14 ] WILLIAMS C K I. Learning with kernels: support vector machines, regularization, optimization, and beyond[J]. Publications of the American Statistical Association,2003,98(462):489-489.
- [ 15 ] RIFKIN R, YEO G, POGGIO T. Regularized least-squares classification[EB/OL]. [2018-08-25]. <http://cbcl.mit.edu/publications/ps/rlsc.pdf>.
- [ 16 ] JIANG Min, SHEN Jianyu, KONG Jun, et al. Robust visual tracking based on kernelized correlation filters[C]//Proceedings of IEEE International Conference on Information and Automation. Washington D. C., USA: IEEE Press,2017:110-115.
- [ 17 ] BECK A, TEBoulLE M. A fast iterative shrinkage-thresholding algorithm for linear inverse problems[J]. SIAM Journal on Imaging Sciences, 2009, 2 ( 1 ) : 183-202.
- [ 18 ] DANELLJAN M, H GER G, KHAN F S, et al. Accurate scale estimation for robust visual tracking [ C ]//Proceedings of British Machine Vision Conference. [ S. l. ]:BMVA Press, 2014:1-11.
- [ 19 ] WU Yi, LIM J, YANG Minghuan. Online object tracking: a benchmark [ C ]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2013:2411-2418.
- [ 20 ] BIBI A, GHANEM B. Multi-template scale-adaptive kernelized correlation filters[C]//Proceedings of IEEE International Conference on Computer Vision. Washington D. C., USA: IEEE Press,2015:613-620.