

基于亚像素修正的模板漂移抑制方法

杨奇龙^{1,2}, 马晓燠^{2,3*}, 刘爽^{1**}, 游双慧², 李成平²¹电子科技大学光电科学与工程学院, 四川 成都 611731;²重庆连芯光电技术研究院有限公司, 重庆 400021;³四川文理学院成都研究院, 四川 成都 635000

摘要 为应对核相关滤波算法长时跟踪模板漂移问题,提出一种基于亚像素修正的模板漂移抑制方法。首先,根据跟踪过程中常用的核相关滤波快速算法,分析模板更新时引入的跟踪误差;其次,分析模板更新时引入的采样误差对模板漂移的影响;然后,通过模板频域的相移修正模板更新引入的采样误差;最后,通过实验证明模板亚像素修正对长时跟踪模板漂移的抑制效果。实验结果表明:相同跟踪条件下,模板亚像素修正方法引入的模板漂移量是整像素模板更新的1/4,模板亚像素修正用于相关滤波算法和核相关滤波算法时对模板漂移同样有良好的抑制效果,并且模板亚像素修正只需耗费较小计算量就可以对模板漂移进行有效抑制。

关键词 图像处理; 核相关滤波; 跟踪算法; 模板漂移; 亚像素

中图分类号 TP391 文献标志码 A

DOI: 10.3788/LOP221228

Template Drift Suppression Method Based on Sub Pixel Correction

Yang Qilong^{1,2}, Ma Xiaoyu^{2,3*}, Liu Shuang^{1**}, You Shuanghui², Li Chengping²¹School of Optoelectronic Science and Engineering, University of Electronic Science and Technology of China, Chengdu 611731, Sichuan, China;²Chongqing Co-Core Optoelectronics Technology Research Institute Co., Ltd., Chongqing 400021, China;³Chengdu Research Institute, Sichuan University of Arts and Sciences, Chengdu 635000, Sichuan, China

Abstract A template drift suppression method based on sub pixel correction is proposed to address the issue of long-term template drift tracking by the kernel correlation filtering algorithm. First, according to the fast algorithm of kernel correlation filtering commonly used in the tracking process, the tracking error introduced by template updating was analyzed. Second, the influence of sampling error on template drift was analyzed. Then, the sampling error introduced by the template update was corrected by shifting the phase of the template frequency domain. Finally, the inhibition effect of the template sub pixel correction on the long-term tracking template drift was verified through experiments. The experimental results show that under the same tracking conditions, the template drift introduced using the template sub pixel correction method is reduced to only 1/4 of that of the whole pixel template update. Therefore, the template sub pixel correction quite effectively suppresses the template drift with a small amount of computation when used in the correlation filtering and kernel correlation filtering algorithms.

Key words image processing; kernel correlation filtering; tracking algorithm; template drift; sub-pixel

1 引言

相关滤波器(CF)因计算简单、跟踪效果好,在视觉跟踪领域大放异彩。2010年,Bolme等^[1]在视觉跟踪中首次应用相关滤波的思想,提出最小误差平方和滤波器(MOSSE)。两年后,Henriques等^[2]以MOSSE

为基础引入核循环矩阵采样以及核相关特征,提出核循环矩阵跟踪(CSK)算法,该算法对目标识别能力大大提升,但在复杂场景下的跟踪效果较差。之后,Henriques等^[3]以CSK为框架引入方向梯度直方图(HOG)特征,改进后的算法为核相关滤波器(KCF),KCF利用多通道目标特征使跟踪滤波器适应能力更

收稿日期: 2022-04-07; 修回日期: 2022-05-06; 录用日期: 2022-07-05; 网络首发日期: 2022-07-15

基金项目: 重庆市自然科学基金(cstc2020jcyj-jqX0019)

通信作者: *maxiaoyu001@126.com; **shuangliu@uestc.edu.cn

强,此后的诸多相关滤波算法都以此为框架。但是,目前的跟踪算法仍无法长时间、稳定地对移动目标进行跟踪。Gong等^[4]认为其问题在于目标的形态和尺度变化、目标被遮挡以及目标图像模糊等都会导致跟踪过程中模板漂移、失效。

在目标跟踪过程中,实时更新模板信息^[5-6]是必须的。为抑制模板漂移问题,目前思路大致有两种:1)选择性地对有效外观模型进行更新。Wu等^[7]成功运用粒子滤波实现自适应模板更新和目标跟踪。Nakhmani等^[8]运用粒子滤波选择匹配度高的子模板更新,可有效防止模板漂移。Fan等^[9]通过在模板上加入权重系数的方法来选择有效的前景信息。Supancic等^[10]以自步学习选择目标可靠观测模型用于模板更新。高赞等^[11]以置信度为权重的方式自适应进行特征融合从而更新模板,防止漂移。2)引入重检测器修正模板漂移。Matthews等^[12]提出的被动漂移校正(PDC)使用初始模板修正跟踪过程中的模板漂移。Ma等^[13]在跟踪的同时训练在线分类器用于目标的重检测,实现对目标的长时目标跟踪(LCT)。Li等^[14]在相关滤波算法结构上加入一个基于支持向量机的检测器用于重检测。Raju等^[15]利用时间信息引导跟踪器和检测器的训练,当目标丢失时以自适应阈值对跟踪器初始化。王依婷^[16]基于互补模型和重检测的方式实现自适应长时跟踪。刘教民等^[17]、陈丽萍等^[18]以及马珺等^[19]结合自适应学习率和目标重检测,实现长时跟踪。李福进等^[20]采用颜色直方图模型结合自适应学习率抑制模板漂移。刘宗达等^[21]在尺度估计跟踪器(DSST)算法上加入自适应学习率和重检测机制,实现长时稳定跟踪。Voigtlaender等^[22]将重检测机制应用于Faster-RCNN中,于2020年在VOT2018-LongTerm数据集中取得优异成绩,其缺点在于跟踪速度太慢。

在跟踪过程中除了目标遮挡、形变等问题引起的模板漂移外,目标的数字采样引入的采样累积误差也是引起模板漂移的原因之一。目标在空间中的运动是连续的,但经过数字采样后变成离散的数字信号。现有跟踪算法模板都是以像素为单位进行更新的,在更新过程中图片中心与目标物实际中心存在亚像素的偏差,这种偏差会随着模板更新累积到新的模板,导致模板漂移量随着更新次数的增加而增加。目前,为抑制模板漂移提出的重监测和高置信度响应帧更新方法,在模板更新时都会引入亚像素偏差造成模板漂移,进而导致跟踪精度降低或跟踪任务失败。目前,对跟踪精度的要求越来越高,从原理上分析这种亚像素模板漂移并讨论解决方案是有必要的。针对现有跟踪算法模板漂移问题,本文从跟踪算法原理上推导模板更新引入的亚像素模板漂移量,提出一种从模板频域的相位引入修正以及模板亚像素修正方法,抑制模板更新时引入的亚像素采样误差。

2 核相关滤波算法原理与工程实现

2.1 核相关滤波算法原理

相关滤波跟踪需要针对观测目标设计一个滤波模板,通过模板在待检测图像不同位置进行互相关运算,寻找最大响应的坐标,该坐标即为待检测图像中的目标位置。核相关滤波算法对待检测图像进行虚拟循环采样,以获得足够多的样本用于分类器训练。而实际计算过程中,用傅里叶域计算替代复杂的循环矩阵求解过程,通过核函数将跟踪问题映射为高维空间中线性可分问题,从而提高算法对前景和背景的判别能力。核相关滤波中核空间的系数矩阵 $\hat{\alpha}$ ^[3]可表示为

$$\hat{\alpha} = \frac{\hat{y}}{\hat{k}_{xx} + \lambda}, \quad (1)$$

式中: x 为训练样本; y 为样本标签; k_{xx} 是 x 的自相关核函数; λ 为正则项; \wedge 表示该变量经傅里叶变换后的频域。检测样本 z 对应的回归值为

$$\hat{f}(z) = \hat{k}_{xz} \hat{\alpha} = \frac{\hat{k}_{xz} \hat{y}}{\hat{k}_{xx} + \lambda}, \quad (2)$$

式中: k_{xz} 是 x 和 z 的核相关。对 $\hat{f}(z)$ 使用傅里叶逆变换可得到响应矩阵。

2.2 核相关滤波算法工程实现

核函数常采用高斯核函数,可表示为

$$k_{xz} = \exp \left\{ \frac{-1}{\sigma^2} \left\{ \|x\|^2 + \|z\|^2 - 2\mathcal{F}^{-1}[\mathcal{F}(x) \cdot \mathcal{F}^*(z)] \right\} \right\}, \quad (3)$$

式中: σ 为高斯核的标准差; \mathcal{F} 表示傅里叶变换。

为了追求实时性,核函数也可选用线性核:

$$k_{xz} = \langle x, z \rangle, \quad (4)$$

式中: $\langle \cdot, \cdot \rangle$ 表示内积。

标签 y 用二维高斯模板作为响应概率:

$$y(s, t) = g(s, t) = \frac{1}{2\pi\sigma^2} \cdot \exp \left(\frac{-s^2 - t^2}{2 \cdot \sigma^2} \right), \quad (5)$$

$$\hat{y}(u, v) = \exp \left\{ -2\pi^2 \sigma^2 \left[\left(\frac{u}{M} \right)^2 + \left(\frac{v}{N} \right)^2 \right] \right\}, \quad (6)$$

式中: s, t 分别表示高斯模板 y 的空间纵横坐标; u, v 分别表示 \hat{y} 在频域空间的纵横坐标; M, N 分别表示 \hat{y} 矩阵的列数和行数。

针对模型的变化,常通过在线学习的方式更新滤波器和观测模型。在对待检测帧进行预测后,由目标预测位置得到新样本 x ,并训练得到新的 $\hat{\alpha}$,最后利用上一帧的模型参数以一定的学习率生成新模型。

滤波器系数 $\hat{\alpha}$ 和目标观测模型 \hat{x} 为

$$\begin{cases} \hat{\alpha}_n = (1 - \gamma) \hat{\alpha}_{n-1} + \gamma \hat{\alpha} \\ \hat{x}_n = (1 - \gamma) \hat{x}_{n-1} + \gamma \hat{x} \end{cases}, \quad (7)$$

式中: γ 为模板更新学习率。

3 模板亚像素修正方法

3.1 模板更新引入的跟踪误差

核相关滤波算法中的滤波器常通过在线更新的方式以适应目标特征的变化。在算法执行完后,得到新的目标预测位置。以当前帧作为训练样本、以预测位置为中心的二维高斯模板作为标签,并通过循环矩阵的方式生成训练样本。

核相关算法实现过程中,主要算力开销在二维傅里叶变换,模板更新时减少二维傅里叶变换次数就可以减小算力开销。以预测位置为中心截取部分图像作为样本 \mathbf{x} , 则 \mathbf{y} 始终为中心分布的二维高斯模板,更新时 $\hat{\mathbf{y}}$ 不变,简化了从 \mathbf{y} 到 $\hat{\mathbf{y}}$ 的计算过程。

问题在于:目标的移动是连续的,但图像采样是离散化的,以预测位置为中心截取的部分图像中,目标的实际位置与模板中心存在亚像素的偏差,如图 1 所示,虚线框表示目标真实位置,实线框表示下一帧模板的采样位置。

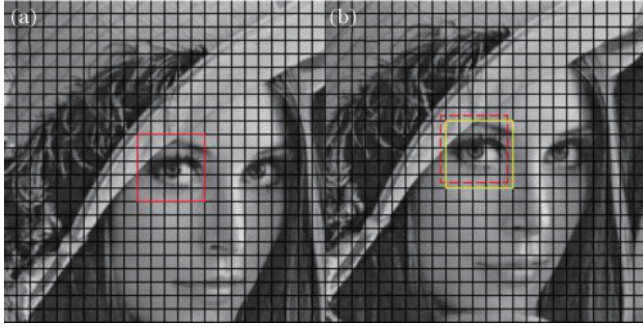


图 1 目标跟踪过程中的亚像素偏差。(a)初始模板位置;
(b)偏移后模板位置与原模板位置

Fig. 1 Sub pixel error in target tracking. (a) Original template position; (b) template position after offset and original template position

3.2 模板更新对模板漂移的影响

设第 n 帧以目标为中心进行离散化采样的目标观测模型真值为 \mathbf{x} , 预测位置为中心像素截取的部分图像 \mathbf{x}' , 截取的图像与观测模型之间的误差为 (ϵ_s, ϵ_t) , 则 \mathbf{x}' 可表示为

$$\mathbf{x}'(s, t) = \mathbf{x}(s - \epsilon_s, t - \epsilon_t). \quad (8)$$

对新截取图像 \mathbf{x}' 进行傅里叶变换后, $\hat{\mathbf{x}}'$ 可表示为

$$\hat{\mathbf{x}}'(u, v) = \hat{\mathbf{x}}(u, v) \exp \left[-2\pi i \left(\frac{u\epsilon_s}{M} + \frac{v\epsilon_t}{N} \right) \right]. \quad (9)$$

将式(9)代入式(1),核函数选用式(4),可得到线性核函数:

$$\hat{\alpha} = \frac{\hat{\mathbf{y}}}{\langle \hat{\mathbf{x}}', \hat{\mathbf{x}}' \rangle + \lambda} = \frac{\hat{\mathbf{y}}}{\| \hat{\mathbf{x}} \|^2 + \lambda}. \quad (10)$$

假设跟踪过程中目标形态不变,则不同帧对应观测模型真值 \mathbf{x} 不变。通过截取图像取样得到的 \mathbf{x}' 与真值 \mathbf{x} 只存在平移误差,将式(9)代入式(7),则在线更新

后的目标观测模型 $\hat{\mathbf{x}}'_n$ 为

$$\begin{aligned} \hat{\mathbf{x}}'_n(u, v) = & \hat{\mathbf{x}}(u, v) \\ & \left\{ (1 - \gamma)^n + \exp \left[-2\pi i \left(\frac{u\epsilon_{s(1)}}{M} + \frac{v\epsilon_{t(1)}}{N} \right) \right] \right\} \times \\ & \gamma (1 - \gamma)^{n-1} + \dots + \\ & \exp \left[-2\pi i \left(\frac{u\epsilon_{s(n-1)}}{M} + \frac{v\epsilon_{t(n-1)}}{N} \right) \right] \times \\ & \gamma (1 - \gamma) + \exp \left[-2\pi i \left(\frac{u\epsilon_{s(n)}}{M} + \frac{v\epsilon_{t(n)}}{N} \right) \right] \gamma \}. \quad (11) \end{aligned}$$

令

$$\begin{aligned} \mathbf{E}_n(u, v) = & \\ & \left\{ (1 - \gamma)^n + \exp \left[-2\pi i \left(\frac{u\epsilon_{s(1)}}{M} + \frac{v\epsilon_{t(1)}}{N} \right) \right] \right\} \times \\ & \gamma (1 - \gamma)^{n-1} + \dots + \\ & \exp \left[-2\pi i \left(\frac{u\epsilon_{s(n-1)}}{M} + \frac{v\epsilon_{t(n-1)}}{N} \right) \right] \times \\ & \gamma (1 - \gamma) + \exp \left[-2\pi i \left(\frac{u\epsilon_{s(n)}}{M} + \frac{v\epsilon_{t(n)}}{N} \right) \right] \gamma \}, \quad (12) \end{aligned}$$

则

$$\hat{\mathbf{x}}'_n(u, v) = \hat{\mathbf{x}}(u, v) \mathbf{E}_n(u, v), \quad (13)$$

$$\begin{aligned} \mathbf{E}_n(u, v) = & (1 - \gamma) \mathbf{E}_{n-1}(u, v) + \\ & \gamma \exp \left[-2\pi i \left(\frac{u\epsilon_{s(n)}}{M} + \frac{v\epsilon_{t(n)}}{N} \right) \right]. \quad (14) \end{aligned}$$

将式(10)代入式(7),在线更新后的滤波器系数 $\hat{\alpha}'_n$ 为

$$\hat{\alpha}'_n = \frac{\hat{\mathbf{y}}}{\| \hat{\mathbf{x}} \|^2 + \lambda}. \quad (15)$$

设目标当前移动量为 (Δ_s, Δ_t) , 待预测帧 \mathbf{z} 可表示为

$$\mathbf{z}(s, t) = \mathbf{x}(s - \Delta_s, t - \Delta_t), \quad (16)$$

则

$$\hat{\mathbf{z}}(u, v) = \hat{\mathbf{x}}(u, v) \exp \left[-2\pi i \left(\frac{u\Delta_s}{M} + \frac{v\Delta_t}{N} \right) \right]. \quad (17)$$

综合式(2)和式(13)~(17)可得:

$$\hat{f}_n(\mathbf{z}) = \mathbf{E}_n^*(u, v) \hat{\mathbf{y}} \exp \left[-2\pi i \left(\frac{u\Delta_s}{M} + \frac{v\Delta_t}{N} \right) \right], \quad (18)$$

式中: $\mathbf{E}_n^*(u, v)$ 为 $\mathbf{E}_n(u, v)$ 的复共轭矩阵。

位置预测误差 $(e_{s(n)}, e_{t(n)})$ 可表示为

$$(e_{s(n)}, e_{t(n)}) = f_n^{-1} \{ \max [f_n(\mathbf{z})] \} - (\Delta_s + \Delta_t). \quad (19)$$

令 $\hat{\mathbf{G}}_n = \mathbf{E}_n^* \hat{\mathbf{y}}$, 则

$$(e_{s(n)}, e_{t(n)}) = \mathbf{G}_n^{-1} \{ \max [\mathbf{G}_n(s, t)] \}. \quad (20)$$

\mathbf{G}_n 可表示为多个高斯核函数的叠加:

$$\begin{aligned} G_n(s, t) = & \mathbf{g}(s, t)(1-\gamma)^n + \mathbf{g}(s - \epsilon_{s(1)}, t - \epsilon_{t(1)}) \times \\ & \gamma(1-\gamma)^{n-1} + \dots + \mathbf{g}(s - \epsilon_{s(n-1)}, t - \epsilon_{t(n-1)}) \times \\ & \gamma(1-\gamma) + \mathbf{g}(s - \epsilon_{s(n)}, t - \epsilon_{t(n)})\gamma. \end{aligned} \quad (21)$$

忽略探测器噪声和拟合等误差,模型观测误差 $(\epsilon_{s(n)}, \epsilon_{t(n)})$ 由上一帧位置预测误差 $(e_{s(n-1)}, e_{t(n-1)})$ 和亚像素采样误差 $(\epsilon_{\text{subs}(n)}, \epsilon_{\text{subt}(n)})$ 组成:

$$(\epsilon_{s(n)}, \epsilon_{t(n)}) = (e_{s(n-1)}, e_{t(n-1)}) + (\epsilon_{\text{subs}(n)}, \epsilon_{\text{subt}(n)}). \quad (22)$$

将 G_{n-1} 近似为以 $(e_{s(n-1)}, e_{t(n-1)})$ 为中心分布的高斯模板,则

$$\begin{aligned} G_n(s, t) = & \mathbf{g}(s - e_{s(n-1)}, t - e_{t(n-1)})(1-\gamma) + \\ & \mathbf{g}(s - \epsilon_{s(n)}, t - \epsilon_{t(n)})\gamma. \end{aligned} \quad (23)$$

从一维高斯模板叠加进行推导,设

$$g(t) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{t^2}{2\sigma^2}\right), \quad (24)$$

将式(24)代入式(23)并求导:

$$\begin{aligned} g_n'(t) = & -(1-\gamma) \frac{t - e_{t(n-1)}}{\sqrt{2\pi}\sigma^3} \exp\left[-\frac{(t - e_{t(n-1)})^2}{2\sigma^2}\right] - \\ & \gamma \frac{t - \epsilon_{t(n)}}{\sqrt{2\pi}\sigma^3} \exp\left[-\frac{(t - \epsilon_{t(n)})^2}{2\sigma^2}\right]. \end{aligned} \quad (25)$$

取 $\sigma > \epsilon_{\text{subt}(n)}$,在 $e_{t(n-1)}$ 和 $\epsilon_{t(n)}$ 之间求解 $g_n'(t) = 0$ 的根,经麦克劳林展开后可表示为

$$\begin{aligned} (1-\gamma)(t - e_{t(n-1)}) \left[1 - \frac{(t - e_{t(n-1)})^2}{2\sigma^2}\right] + \\ \gamma(t - \epsilon_{t(n-1)} - \epsilon_{\text{subt}(n)}) \left[1 - \frac{(t - \epsilon_{t(n-1)} - \epsilon_{\text{subt}(n)})^2}{2\sigma^2}\right] = 0, \end{aligned} \quad (26)$$

化简后由盛金公式求解,取 $e_{t(n-1)}$ 和 $\epsilon_{t(n)}$ 之间的根:

$$t - e_{t(n-1)} = \frac{-b + \sqrt{A} \left(\cos \frac{\theta}{3} - \sqrt{3} \sin \frac{\theta}{3} \right)}{3a}, \quad (27)$$

式中: $\theta = \arccos T$, $T = \frac{2Ab - 3aB}{2A\sqrt{A}}$, $A = b^2 - 3ac$,

$B = bc - 9ad$, $a = 1$, $b = 3\epsilon_{\text{subt}(n)}\gamma$, $c = 3\epsilon_{\text{subt}(n)}^2\gamma - 2\sigma^2$,

$d = 2\epsilon_{\text{subt}(n)}\gamma\sigma^2 - \epsilon_{\text{subt}(n)}^3\gamma$ 。当 $\epsilon_{\text{subt}(n)}$ 和 γ 都小于1且远

小于 σ 时, $T \rightarrow 0$, $\theta \approx \frac{\pi}{2} - T$,经泰勒展开后,有

$\cos \frac{\theta}{3} - \sqrt{3} \sin \frac{\theta}{3} \approx \frac{2T}{3}$,代入式(27)后可得到 $t -$

$e_{t(n-1)}$ 的近似解:

$$\begin{aligned} t - e_{t(n-1)} \approx & -\frac{1}{3} \epsilon_{\text{subt}(n)}\gamma - \\ & \frac{3\epsilon_{\text{subt}(n)}^3\gamma^2 + 3\epsilon_{\text{subt}(n)}^3\gamma - 8\epsilon_{\text{subt}(n)}\gamma\sigma^2}{9\epsilon_{\text{subt}(n)}^2\gamma^2 - 9\epsilon_{\text{subt}(n)}^2\gamma + 6\sigma^2} \approx \epsilon_{\text{subt}(n)}\gamma. \end{aligned} \quad (28)$$

推广到二维高斯叠加,有

$$(e_{s(n)}, e_{t(n)}) \approx (e_{s(n-1)}, e_{t(n-1)}) + \gamma(\epsilon_{\text{subs}(n)}, \epsilon_{\text{subt}(n)}), \quad (29)$$

则

$$(e_{s(n)}, e_{t(n)}) \approx \sum_{i=1}^n \gamma(\epsilon_{\text{subs}(i)}, \epsilon_{\text{subt}(i)}). \quad (30)$$

如果 $\epsilon_{\text{subs}(i)}$ 和 $\epsilon_{\text{subt}(i)}$ 都服从均匀分布 $U(-0.5, 0.5)$,则 $e_{s(n)}$ 的期望为0,标准差为 $\gamma\sqrt{n/12}$ 。经过不断更新,模板可能的漂移范围会越来越大。通过MATLAB仿真误差 $U(-0.5, 0.5)$ 的误差累积结果,并与理论标准差 $\sqrt{n/12}$ 对比,结果如图2所示。

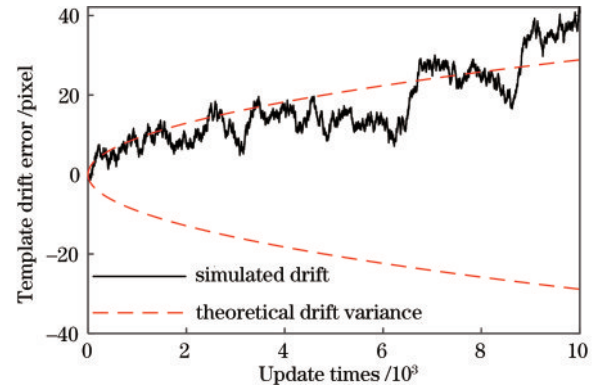


图2 模板漂移与更新次数的关系

Fig. 2 Relationship between template drift and update times

3.3 模板亚像素修正方法

设以第 $n-1$ 帧预测位置为中心进行离散化采样的目标观测模型与以预测位置为中心像素截取的部分图像存在亚像素的位置平移为 $(\epsilon_{\text{subs}(n)}, \epsilon_{\text{subt}(n)})$ 。

可以通过双线性插值的方法修正亚像素位置平移。双线性插值原理如图3所示,为计算亚像素坐标 $(s + \epsilon_{\text{subs}}, t + \epsilon_{\text{subt}})$ 的像素值 $I(s + \epsilon_{\text{subs}}, t + \epsilon_{\text{subt}})$,其中, s, t 为整数, ϵ_{subs} 和 ϵ_{subt} 为大于0小于1的小数。选取被插值像素点邻域的4个像素 $(s, t), (s + 1, t), (s, t + 1), (s + 1, t + 1)$ 所对应的像素值,通过一定的权重计算方法得出亚像素位置像素点数值:

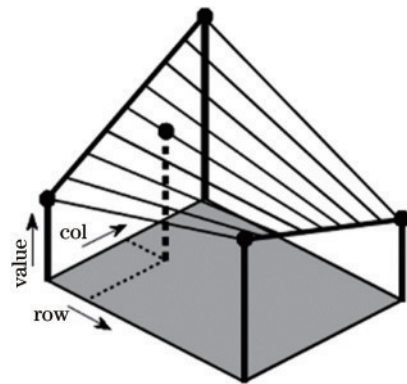


图3 双线性插值示意图

Fig. 3 Schematic diagram of bilinear interpolation

$$I(s + \epsilon_{\text{subs}}, t + \epsilon_{\text{subr}}) = (1 - \epsilon_{\text{subs}})(1 - \epsilon_{\text{subr}})I(s, t) + (1 - \epsilon_{\text{subs}})\epsilon_{\text{subr}}I(s, t + 1) + \epsilon_{\text{subs}}(1 - \epsilon_{\text{subr}})I(s + 1, t) + \epsilon_{\text{subs}}\epsilon_{\text{subr}}I(s + 1, t + 1). \quad (31)$$

双线性插值可实现图片的亚像素平移,但得到亚像素平移修正后的样本之后还需要再进行傅里叶变换才能进行模板更新。算法实现过程中傅里叶变换会消耗大量算力,为减少算法中傅里叶变换的次数可通过对图像频域的相移等效实现图像在空域的平移。

设第 n 帧预测帧为 z_n , 其傅里叶变换为 \hat{z}_n , 经相关滤波预测目标位置为 (δ_s, δ_t) , 则观测模型 $\mathbf{x}(s, t)$ 为

$$\mathbf{x}(s, t) = \mathbf{z}(s + \delta_s, t + \delta_t), \quad (32)$$

频域为

$$\hat{\mathbf{x}}(u, v) = \hat{\mathbf{z}}(u, v) \exp\left[2\pi i \left(\frac{u\delta_s}{M} + \frac{v\delta_t}{N}\right)\right]. \quad (33)$$

(δ_s, δ_t) 可以是整数也可以是小数,通过频域的相移实现了空域的亚像素平移,弥补了传统截图取样引入的亚像素误差,消除了因更新采样引起的模板漂移。

3.4 模板亚像素修正步骤

- 1) 通过式(5)和式(6)生成响应模板;
- 2) 根据观测模型 $\hat{\mathbf{x}}$ 通过式(3)或式(4)实现图像向核空间映射;
- 3) 根据观测模型 $\hat{\mathbf{x}}$ 通过式(1)求系数矩阵 $\hat{\alpha}$;
- 4) 获取预测图像进行傅里叶变换得到 \hat{z} ;
- 5) 通过式(2)得到预测响应 \hat{f} , 进行傅里叶逆变换得到预测目标位置 (δ_s, δ_t) ;
- 6) 通过式(33)对观测模型进行采样亚像素修正,修正后的观测模型代入式(7)进行更新。

4 测试实验与结果

4.1 实验室环境下模板漂移抑制实验

4.1.1 实验设定

主要推导采样过程中引入的亚像素误差对模板漂移的影响,并引入图片频域相移修正因采样引起的亚像素误差。为了验证核相关滤波算法中模板亚像素修正对模板漂移的抑制效果,在目标移动过程中目标的形状、大小和姿态不发生变化,以排除其他因素的影响。

实验场景如图 4 所示,图像采集设备为 1280×1024 的工业相机,像元尺寸为 $4.8 \mu\text{m} \times 4.8 \mu\text{m}$,镜头焦距为 75 mm 。被测量目标为打印的模拟目标,大小约为 $10 \text{ mm} \times 10 \text{ mm}$,距离图像采集设备约 0.75 m ,目标放置于电动平移台上,可垂直于观测方向水平移动,电动平移台的定位精度为 $2 \mu\text{m}$ 。可以通过电动平移台的移动量验证跟踪算法的精度与模板亚像素修正对模板漂移的抑制效果。

4.1.2 实验过程

图像采集设备正对被测电动平移台,采集图像作为初始模板;以随机步长水平移动靶标,每次移动完成

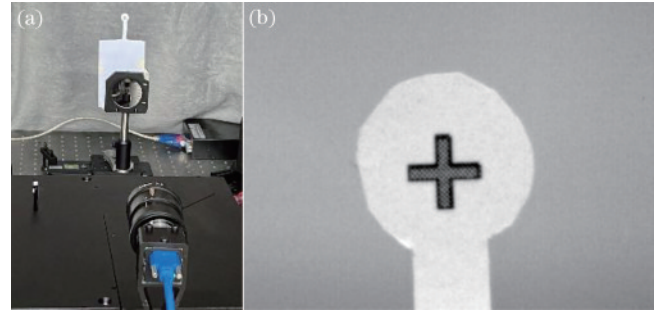


图 4 实验场景与实验图像。(a)实验场景;(b)实验图像
Fig. 4 Experimental scene and image. (a) Experimental scene; (b) experimental image

后存储图像和当前靶标位置;采集 1800 组数据,以初始模板对图像序列进行滤波并记录匹配结果;选定一个算法分别以整像素模板更新和亚像素修正后模板对图像序列进行滤波,并将两种更新方式匹配结果与初始模板匹配结果对比,验证核相关滤波算法中模板亚像素修正对模板漂移的抑制效果。

4.1.3 实验结果

分别以整像素模板更新和模板亚像素修正对图像序列进行相关滤波:以原算法匹配结果减初始模板匹配结果作为校正前模板漂移量,以亚像素模板修正后的算法匹配结果减初始模板匹配结果作为校正后模板漂移量。对 MOSSE 算法模板校正前后漂移误差如图 5 所示。为在较短序列中表现出模板漂移,设置模板学习率为 0.2,从误差图可以看出,经过 1800 帧更新后, MOSSE 算法整像素模板更新造成的模板漂移量为 -21.2 pixel ,远大于 $\gamma\sqrt{n/12}$ 。其主要原因在于模板更新除亚像素误差外还引入了镜头畸变、相机噪声、峰值拟合等其他误差。采用亚像素模板更新方法修正后,经过 1800 帧更新后的模板漂移量为 -5.6 pixel 。可以看出,模板亚像素修正对 MOSSE 算法模板漂移有良好的抑制效果。

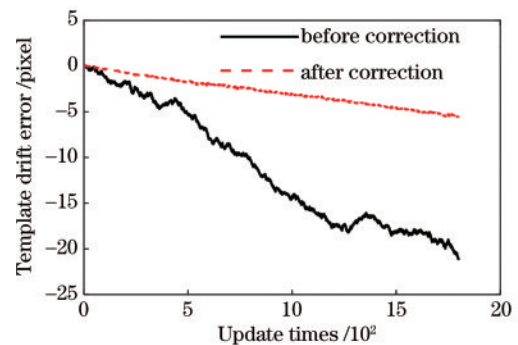


图 5 不同更新方法引起的模板漂移量(MOSSE)
Fig. 5 Comparison of template drift caused by different updating methods (MOSSE)

亚像素模板修正方法用于 CSK 和 KCF 算法同样对模板漂移有良好的抑制效果,如图 6 和图 7 所示。经

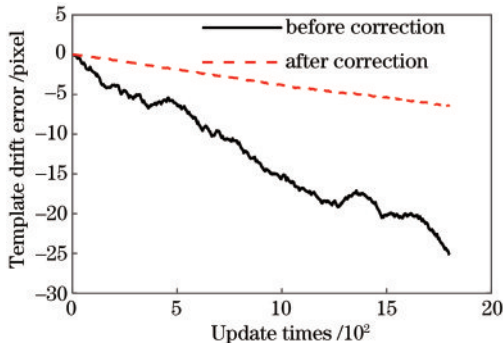


图 6 不同更新方法引起的模板漂移量(CSK)

Fig. 6 Comparison of template drift caused by different updating methods (CSK)

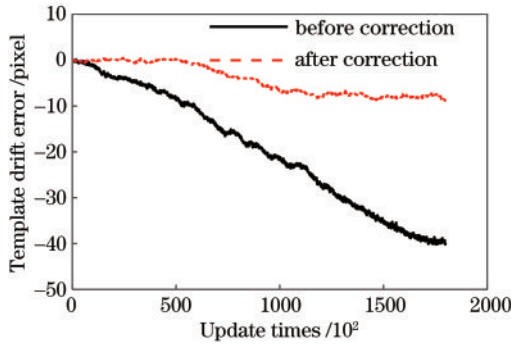


图 7 不同更新方法引起的模板漂移量(KCF)

Fig. 7 Comparison of template drift caused by different updating methods (KCF)

过 1800 帧更新后,CSK 算法整像素模板更新造成的模板漂移量为 -25.2 pixel,采用亚像素模板更新方法修正后,经过 1800 帧更新后的模板漂移量为 -6.4 pixel。经过 1800 帧更新后,KCF 算法整像素模板更新造成的模板漂移量为 -39.9 pixel,采用亚像素模板更新方法修正后,经过 1800 帧更新后的模板漂移量为 -8.72 pixel。

4.2 主流跟踪算法校正结果

为了验证所提算法的有效性,将其在 MATLABR 2018b 软件中进行仿真模拟,实验环境为 Intel(R) Core(TM) i7-1165G7 @2.80 GHz,选取公开数据集 OTB-2015 中 100 个视频序列中的 7 个视频序列进行测试,7 个视频分别为“dog1”“fish”“mhyang”“mountain bike”“sylvestre”“trellis”“walking”。选取 CSK、KCF、fDSST 等 3 种当下主流实时跟踪算法进行亚像素校正,并分析校正前后的跟踪性能。

采用跟踪精确度(DP)和成功率(OP)作为衡量亚像素模板校正效果的标准,采用跟踪帧率(FPS)来评价亚像素模板校正算法实时性。其中:跟踪精确度表示中心位置误差小于给定阈值的帧数与全部视频序列总帧数的比值,本实验阈值设为 20 pixel;成功率表示算法目标框重合率大于某个设定阈值的帧数占总体视频帧数的比值,本实验阈值为 0.5。

CSK、KCF、fDSST 经亚像素模板校正后分别记为 CSK_Cor、KCF_Cor、fDSST_Cor。实验中设 CSK、KCF 和 fDSST 的默认学习率分别为 0.075、0.02、0.025,CSK_Cor、KCF_Cor、fDSST_Cor 的学习率同样为 0.075、0.02、0.025。精度曲线图和成功率曲线图如图 8 和图 9 所示,从精度图和成功率图可以看出,经亚像素模板校正后 3 种算法的跟踪精度和成功率均优于校正前的算法。但由于数据集中视频序列长度都在 1000 帧左右,且学习率不高,根据第 3.2 节中结论可推算出因亚像素采样偏差引起的模板漂移约为 0.75 pixel,所以精度提升不大。

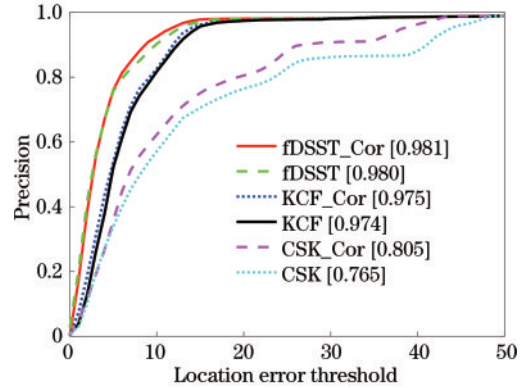


图 8 默认学习率下几种算法的跟踪精度曲线图

Fig. 8 Precision curves of different methods with default learn rate

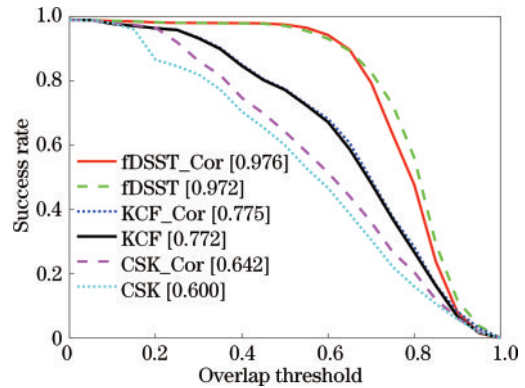


图 9 默认学习率下几种算法的跟踪成功率曲线图

Fig. 9 Success rate curves of different methods with default learn rate

为在较短的视频序列里表现长时跟踪才会出现的模板漂移,分别将 CSK、KCF 和 fDSST 这 3 种算法的学习率增加为 0.09、0.05、0.05,CSK_Cor、KCF_Cor、fDSST_Cor 的学习率同样设为 0.09、0.05、0.05。大学习率条件下的精度曲线图和成功率曲线图如图 10 和图 11 所示,从新的精度图和成功率图可以看出,经亚像素模板校正后 3 种算法的跟踪精度和成功率均优于校正前的算法,且大学习率实验中经亚像素模板校正后 3 种算法的跟踪精度和成功率提升明显。

对算法进行亚像素校正后,除跟踪精度提升外还可以提升算法的跟踪稳定性,具体如图 12 所示。在

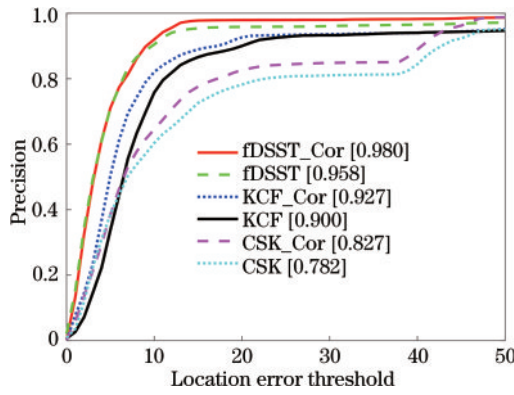


图 10 大学习率条件下几种算法的跟踪精度曲线图
Fig. 10 Precision curves of different methods with large learn rate

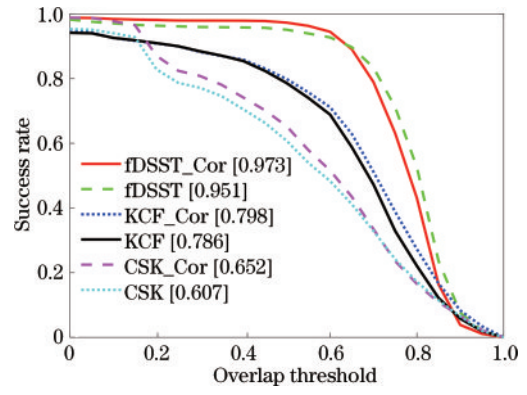


图 11 大学习率条件下几种算法的跟踪成功率曲线图
Fig. 11 Success rate curves of different methods with large learn rate

“dog1”视频序列中 CSK 和 KCF 因目标尺度变化会引起漂移但 CSK_Cor 和 KCF_Cor 仍对目标中心位置保持跟踪;“fish”视频序列中因光照变化太快 CSK 和 CSK_Cor 均出现漂移;“mhyang”视频序列中 KCF 和 fDSST 因光照变化引起漂移但 KCF_Cor 和 fDSST_Cor 仍对目标中心位置保持跟踪;“mountain

bike”视频序列中 6 种算法均表现良好;“sylvester”视频序列中因背景杂乱 fDSST 跟踪失败而 fDSST_Cor 仍对目标中心位置保持跟踪;“trellis”视频序列中因光照变化太快 CSK 和 CSK_Cor 均出现漂移;“walking”视频序列中 KCF 因目标分辨率低以及尺度变化引起漂移但 KCF_Cor 仍对目标保持跟踪。

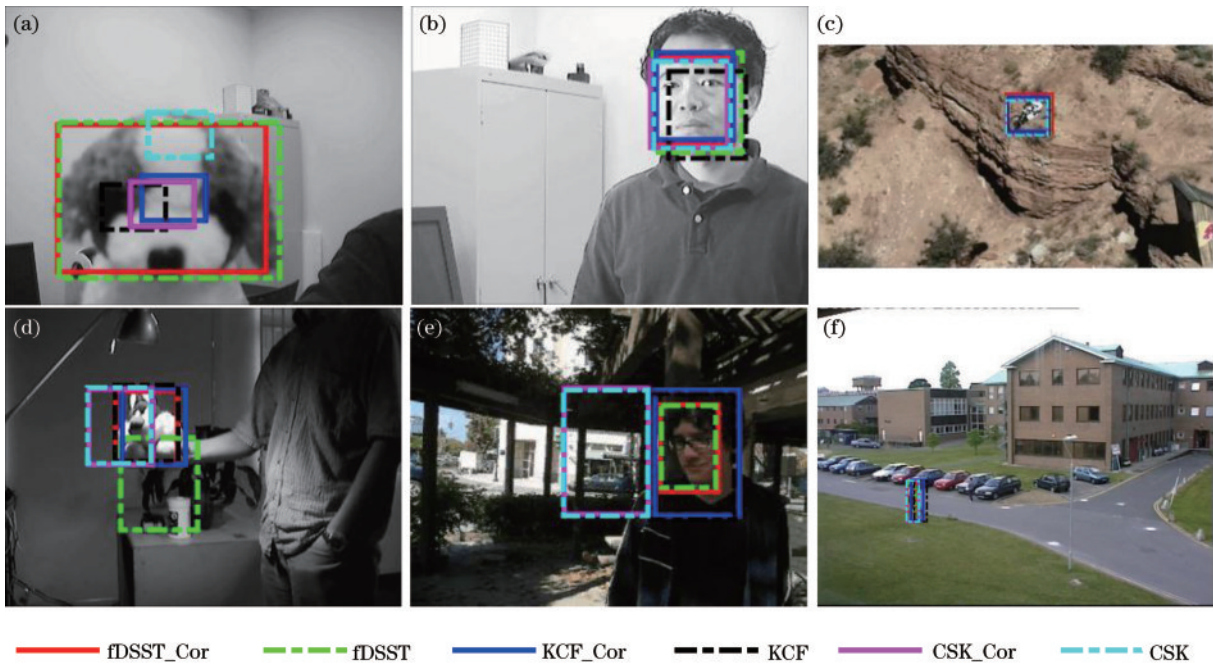


图 12 6 种算法对不同视频序列的跟踪结果。(a) Dog1; (b) mhyang; (c) mountainbike; (d) sylvester; (e) trellis; (f) walking
Fig. 12 Tracking results of six algorithms for different video sequences. (a) Dog1; (b) mhyang; (c) mountainbike; (d) sylvester; (e) trellis; (f) walking

分别采用 6 种算法对 7 个视频序列的跟踪速度进行统计,结果如表 1 所示:fDSST 平均跟踪速度为 111.3 frame/s, fDSST_Cor 平均跟踪速度为 87.6 frame/s, fDSST_Cor 相比 fDSST 性能平均降低 22.84%; KCF 平均跟踪速度为 191.2 frame/s, KCF_Cor 平均跟踪速度为 182.6 frame/s, KCF_Cor

相比 KCF 性能平均降低 4.45%; CSK 平均跟踪速度为 329.5 frame/s, CSK_Cor 平均跟踪速度为 271.6 frame/s, CSK_Cor 相比 CSK 性能平均降低 17.13%。通过对比发现, KCF 进行亚像素校正后跟踪性能只下降了 4.45%, 主要是因为 KCF 算法的 HOG 特征的空间维度较小, 进行傅里叶变换耗时较少。

表 1 6 种算法对不同视频序列的跟踪速度

Table 1 Tracking speed of six algorithms for different video sequences

Video sequence	fDSST / (frame·s ⁻¹)	fDSST_Cor/ (frame·s ⁻¹)	Performance lost /%	KCF/ (frame·s ⁻¹)	KCF_Cor/ (frame·s ⁻¹)	Performance lost /%	CSK/ (frame·s ⁻¹)	CSK_Cor/ (frame·s ⁻¹)	Performance lost /%
dog1	141.4	125.0	11.60	300.6	289.7	3.63	505.0	419.6	16.91
fish	98.2	83.0	15.48	128.3	122.5	4.52	332.9	265.7	20.19
mhyang	109.5	90.3	17.53	139.2	133.1	4.38	308.8	250.0	19.04
mountainbike	97.1	65.0	33.06	175.2	167.2	4.57	305.1	245.7	19.47
sylvester	131.9	109.8	16.76	195.5	185.6	5.06	226.2	204.7	9.50
trellis	72.4	42.7	41.02	85.5	82.2	3.86	198.9	167.7	15.69
walking	128.6	97.1	24.45	314.1	297.9	5.16	429.7	347.7	19.10
Average	111.3	87.6	22.84	191.2	182.6	4.45	329.5	271.6	17.13

5 结 论

针对现有跟踪算法模板漂移问题,从跟踪算法原理上推导出,模板更新引入的亚像素模板漂移量与 \sqrt{n} 成正比(n 为模板更新帧数)。提出从模板频域的相位引入修正的方法,最少只需要进行一次欧拉公式计算和频域矩阵即可对模板进行修正,在普通设备就能以少量计算量抑制模板更新时引入的亚像素采样误差。实验结果表明:相同跟踪条件下模板亚像素修正方法引入的模板漂移量是整像素模板更新的1/4,模板亚像素修正方法用于相关滤波算法和核相关滤波算法时对模板漂移同样有良好的抑制效果。将所提方法分别应用于CSK、KCF和fDSST在TOB-2015中的7个视频进行了跟踪精度、成功率和跟踪速度测试,实验结果表明,该方法对3种算法的跟踪精度和成功率均有不同程度的提升,在对KCF进行校正时只降低了原来算法4.45%的跟踪速度。现有算法解决模板漂移时主要针对目标的形态和尺度变化、目标遮挡以及目标图像模糊等问题进行改进,但采样引起的亚像素模板漂移在长时间跟踪和大学学习率跟踪任务中尤为突出,而现有的模板更新方法并未考虑。随着对跟踪算法的精度和稳定性要求的不断提高,所提亚像素模板漂移误差和抑制方法或为跟踪算法优化提供一个新的方向。

参 考 文 献

- [1] Bolme D S, Beveridge J R, Draper B A, et al. Visual object tracking using adaptive correlation filters[C]//2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, June 13-18, 2010, San Francisco, CA, USA. New York: IEEE Press, 2010: 2544-2550.
- [2] Henriques J F, Caseiro R, Martins P, et al. Exploiting the circulant structure of tracking-by-detection with kernels[M]//Fitzgibbon A, Lazebnik S, Perona P, et al. Computer vision-ECCV 2012. Lecture notes in computer science. Heidelberg: Springer, 2012, 7575: 702-715.
- [3] 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.
- [4] Gong H F, Sim J, Likhachev M, et al. Multi-hypothesis motion planning for visual object tracking[C]//2011 International Conference on Computer Vision, November 6-13, 2011, Barcelona. New York: IEEE Press, 2011: 619-626.
- [5] 姜学军, 万晓东. 模版漂移纠正的目标跟踪算法研究[J]. 现代计算机, 2013(24): 25-29.
Jiang X J, Wan X D. Template object tracking algorithm with weighted drift correction[J]. Modern Computer, 2013(24): 25-29.
- [6] 谢郭蓉, 曲毅, 蒋榕圻. 基于抗遮挡目标模型的跟踪算法综述[J]. 激光与光电子学进展, 2022, 59(8): 0815001.
Xie G R, Qu Y, Jiang R Q. Tracking algorithms based on antiocclusion object models[J]. Laser & Optoelectronics Progress, 2022, 59(8): 0815001.
- [7] Wu P P, Wang J L, Zhang Q. Covariance tracking algorithm based on particle filter and adaptive template update[C]//2010 2nd International Conference on Education Technology and Computer, June 22-24, 2010, Shanghai, China. New York: IEEE Press, 2010: 199-202.
- [8] Nakhmani A, Tannenbaum A. Particle filtering with region-based matching for tracking of partially occluded and scaled targets[J]. SIAM Journal on Imaging Sciences, 2011, 4(1): 220-242.
- [9] Fan B J, Du Y K, Zhu L L, et al. A robust template tracking algorithm with weighted active drift correction[J]. Pattern Recognition Letters, 2011, 32(9): 1317-1327.
- [10] Supancic J S, Ramanan D. Self-paced learning for long-term tracking[C]//2013 IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2013, Portland, OR, USA. New York: IEEE Press, 2013: 2379-2386.
- [11] 高赞, 赵江珊, 罗久桓, 等. 采用响应图置信区域自适应特征融合的相关滤波跟踪[J]. 光学精密工程, 2019, 27(5): 1178-1187.
Gao Y, Zhao J S, Luo J H, et al. Adaptive feature fusion with the confidence region of a response map as a correlation filter tracker[J]. Optics and Precision Engineering, 2019, 27(5): 1178-1187.
- [12] Matthews L, Ishikawa T, Baker S. The template update

- problem[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 26(6): 810-815.
- [13] Ma C, Yang X K, Zhang C Y, et al. Long-term correlation tracking[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition, June 7-12, 2015, Boston, MA, USA. New York: IEEE Press, 2015: 5388-5396.
- [14] Li T, Zhao S Y, Meng Q H, et al. A stable long-term object tracking method with re-detection strategy[J]. Pattern Recognition Letters, 2019, 127: 119-127.
- [15] Raju P M, Mishra D, Gorthi R K S S. Detection based long term tracking in correlation filter trackers[J]. Pattern Recognition Letters, 2019, 122: 79-85.
- [16] 王依婷. 基于相关滤波的目标跟踪算法研究[D]. 西安: 西北大学, 2020.
Wang Y T. Research on target tracking algorithm based on correlation filtering[D]. Xi'an: Northwest University, 2020.
- [17] 刘教民, 郭剑威, 师硕. 自适应模板更新和目标重定位的相关滤波器跟踪[J]. 光学精密工程, 2018, 26(8): 2100-2111.
Liu J M, Guo J W, Shi S. Correlation filter tracking based on adaptive learning rate and location refiner[J]. Optics and Precision Engineering, 2018, 26(8): 2100-2111.
- [18] 陈丽萍, 王铭羽, 杨文柱, 等. 基于改进核相关滤波的长时目标跟踪算法[J]. 郑州大学学报(工学版), 2020, 41(3): 32-36.
Chen L P, Wang M Y, Yang W Z, et al. An improved kernelized correlation filter for long-term target tracking[J]. Journal of Zhengzhou University (Engineering Science), 2020, 41(3): 32-36.
- [19] 马珺, 王昱皓. 结合自适应更新策略和再检测技术的跟踪算法[J]. 计算机工程与应用, 2021, 57(9): 217-224.
Ma J, Wang Y H. Object tracking algorithm based on adaptive update strategy and re-detection technology[J]. Computer Engineering and Applications, 2021, 57(9): 217-224.
- [20] 李福进, 刘慧慧, 任红格, 等. 高置信度的尺度自适应核相关跟踪方法[J]. 激光与光电子学进展, 2021, 58(8): 0815004.
Li F J, Liu H H, Ren H G, et al. Scale adaptive kernel correlation tracking method with high confidence[J]. Laser & Optoelectronics Progress, 2021, 58(8): 0815004.
- [21] 刘宗达, 董立泉, 赵跃进, 等. 视频中快速运动目标的自适应模型跟踪算法[J]. 光学学报, 2021, 41(18): 1815001.
Liu Z D, Dong L Q, Zhao Y J, et al. Adaptive model tracking algorithm for fast-moving targets in video[J]. Acta Optica Sinica, 2021, 41(18): 1815001.
- [22] Voigtlaender P, Luiten J, Torr P H S, et al. Siam R-CNN: visual tracking by re-detection[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 6577-6587.