E-mail: jig@aircas.ac.cn
Website: www.cjig.cn
Tel: 010-58887035
中国图象图形学报
JOURNAL OF IMAGE AND GRAPHICS
©中国图象图形学报版权所有

中图法分类号:TP391　　**文献标识码**: A　　**文章编号**: 1006-8961(2023)10-3092-15

论文引用格式:Wang F S, Yin S S, He B and Sun F M. 2023. A Gaussian mask-based correlation filter tracking algorithm. Journal of Image and Graphics, 28(10):3092-3106(王法胜, 尹双双, 贺冰, 孙福明. 2023. 引入高斯掩膜的相关滤波目标跟踪算法. 中国图象图形学报, 28(10):3092-3106)[DOI:10. 11834/jig. 220856]

引入高斯掩膜的相关滤波目标跟踪算法

王法胜,尹双双,贺冰,孙福明*

大连民族大学信息与通信工程学院,大连　116600

摘　要: **目的**　由于相关滤波(correlation filter,CF)跟踪算法使用循环移位操作增加训练样本,不可避免会引起边界效应问题。研究者大多采用余弦窗来抑制边界效应,但余弦窗的引入会导致样本污染,降低了算法的性能。为解决该问题,提出了嵌入高斯形状掩膜的相关滤波跟踪算法。**方法**　在空间正则化相关滤波跟踪框架中嵌入高斯形状掩膜,对靠近目标中心和搜索区域边缘的样本重新分配权重,提高中心样本的重要性,同时降低远离中心的边缘样本的重要性,增加中心样本的响应,从而降低样本污染的影响,增强滤波器的判别能力。建立了高斯掩膜相关滤波跟踪算法的目标公式,然后使用交替方向乘子法(alternating direction of multiplier method,ADMM)求解滤波器及空间权重的闭合解。**结果**　为评估所提算法的跟踪性能,在OTB2013(online tracking benchmark)、TC128(temple color)、UAV123(unmmaned aerial vehicle)及Got-10k(general object tracking)等多个基准数据集上进行了大量实验,并与多个先进的相关滤波跟踪算法对比。结果表明,在OTB2013数据集上精度和成功率分别为90. 2%和65. 2%,其中精度比基准算法提高0. 5%;在TC128数据集上精度和成功率分别为77. 9%和57. 7%,其中成功率提高0. 4%;UAV123数据集上的两个指标数据分别为74. 1%和50. 8%,精度提高0. 3%;在Got-10k数据集上成功率提高了0. 2%。**结论**　与其他相关滤波跟踪算法相比,本文算法在各数据集上的精度和成功率表现出较强的竞争力,显著提高了基准算法的跟踪性能。

关键词: 目标跟踪;相关滤波(CF);边界效应;样本污染;高斯形状掩膜

A Gaussian mask-based correlation filter tracking algorithm

Wang Fasheng, Yin Shuangshuang, He Bing, Sun Fuming*

School of Information and Communication Engineering, Dalian Minzu University, Dalian　116600, China

Abstract: **Objective** Object tracking has been concerned more as such a domain in relevance to computer vision. In the past decade, correlation filter has been recognized for object tracking due to its high accuracy and speed. For traditional correlation filter tracking algorithms, training sample set is generated using circulant shift operation on the underlying image patches, and the circulant matrix is used to compute in the Fourier domain as well, which can improve the learning efficiency further. However, due to Fourier transform is used periodically, the boundary of the image is not involved in computing. Consequently, the boundary of the object-tracking sample is discontinuous, resulting in the annoying boundary effect. The error and inquality of samples can be generated, which affects the filter ability severely. Two sorts of popular

收稿日期:2022-09-06;**修回日期**:2022-11-22;**预印本日期**:2022-11-29

* **通信作者**:孙福明　sunfuming@dlnu.edu.cn

基金项目:国家自然科学基金项目(61972068,61976042);兴辽英才计划项目(XLYC2007023);辽宁省高等学校创新人才支持计划项目(LR2019020)

Supported by: National Natural Science Foundation of China(61972068, 61976042); Liaoning Revitalization Talents Program (XLYC2007023); Innovative Talents Program for Liaoning Universities(LR2019020)

solutions are employed to resolve the boundary effects currently. The first one is oriented to draw real negative samples from the sampling area. But, all of the negative and positive samples are assigned equal weights, which cannot well suppress the boundary effect. The other one is focused on an integrated spatial regularization term relevant to the correlation filter framework. A defined spatial weight map is used to deal with boundary area-near spatial weights of the samples. However, spatial regularization weight map is defined in the first frame, and such related invariant can be kept during tracking, which is not melted into the variation of different object. Both of the two sorts of solutions are required to suppress boundary effects using cosine window. The introduction of cosine window is linked to the problem of sample contamination simultaneously. The generated samples are distorted by redundant pixels, for which the tracking performance is eroded. To resolve this problem, we develop a Gaussian mask-based correlation filter tracking algorithm. **Method** Gaussian shape mask is embedded into the spatially regularized correlation filter framework. It redefines the weight of the samples close to the object center and those next to the boundary of the sampling region, which strengthens the weight of the samples the object center and weakens the weight of the samples far away from the object center. Thus, the response of the samples close to the object center could be enhanced while those boundary area-near samples are suppressed. To improve the tracking performance, its related strategy can be implemented to suppress the boundary effect and shrink cosine window-caused sample contamination problem. The objective function of the proposed Gaussian mask correlation filter is configured out, and to obtain closed-form solutions of the filter and spatial weights, it is then solved using alternating direction of multiplier method. To deal with scale estimation problem, two sort of correlation filters are trained in terms of location and scale filters. The location filter is trained using histogram of gradient (HOG) feature and Visual Geometry Group (VGG) network-derived deep features, while the scale filter is trained using five-scale HOG features. The former one is mainly used for estimating the object location, while the latter one is used to determine the best scale of the target object. **Result** To evaluate the proposed Gaussian mask correlation filter, extensive experiments are carried out on several tracking benchmark datasets relevant to OTB2013, TC128, UAV123 and Got-10k, and comparative analysis is linked to such popular correlation filter based tracking methods like LADCF, ASRCF, RHCF and ECO. We use the general one-pass evaluation standard to evaluate all the tracking methods. For the evaluation metrics, the precision score and success rate are used and all the tracking methods are ranked according to the precision score in the precision plots and area under curve score in the success plots. The experimental results demonstrate that each of the OTB2013 dataset-relevant precision and success score is reached to 90. 2% and 65. 2%, in which the precision score is optimized by 0. 5% over the baseline tracking method. For TC128 dataset, the values of the two metrics are 77. 9% and 57. 7%, where the success score is improved by 0. 4 over the baseline tracking method. For UAV123 dataset, each of the two sorts of scores can be reached to 74. 1% and 50. 8%, where the precision score is improved by 0. 3% compared to the baseline tracking method. For Got-10k dataset, the success rate is improved by 0. 2% compared to the baseline tracking method. The qualitative results illustrate that our proposed method can be used to deal with such challenging issues like deformation, occlusion, out of view and fast motion. However, the limitation shows that computational cost of the proposed method is inevitable due to the introduction of the Gaussian mask. In addition, VGG network based deep features can increase the computational cost as well. For Got-10k dataset, the frame rate is about 5. 2 frames per second, while the baseline tracking method is 5. 43 frames per second, which realizes 0. 23 frames per second higher. **Conclusion** The Gaussian mask is melted into the correlation filter framework, which can suppress the boundary effect and resolve the sample contamination problem. The sample weights are redistributed to highlight the samples near the object center, while the weight of the samples far away from the object center are suppressed. Extensive experimental results demonstrate that the proposed strategy is effective to some extent. Compared to other related correlation filter based trackers, our proposed method has its potentials on four public benchmark datasets in terms of tracking precision and success rate, which can improves the tracking performance of the baseline tracking method further. Guassian mask can also be extended to other related correlation filter based tracking framework potentially. The introduction of the Gaussian mask can decrease the tracking speed. Our future work will focus on this problem.

Key words: object tracking; correlation filter (CF); boundary effect; sample contamination; Gaussian shape mask

0 引 言

目标跟踪在现实世界中有着广泛的应用，已经成为计算机视觉领域的一个重要研究方向（Abbass等，2021）。近年来，目标跟踪技术在国内外学者的努力下取得了长足的进步，但在复杂场景中，由于缺失目标信息，导致传统的跟踪方法很难同时满足跟踪精度和实时性两个方面的要求。相关滤波算法（correlation filter，CF）由Bolme等人（2010）首次引入目标跟踪领域后，在多种挑战场景中表现出良好的精确度和实时性，成为一种解决目标跟踪问题的主流方法（强壮和石繁槐，2020）。在经典的相关滤波跟踪算法中，训练集是利用基础图像的循环移位来生成正负样本，并且利用循环矩阵在傅里叶域中进行计算，显著提高了学习效率。但由于傅里叶变换是周期性的，并没有考虑到图像的边界，使得跟踪物体样本的边界不连续，引发了边界效应（Liu等，2021）。边界效应会导致错误样本的产生，降低样本的质量，从而严重影响滤波器的判别能力。

为解决边界效应问题，早期的相关滤波算法使用余弦窗与目标区域相乘使目标边缘像素接近于零。由于滤波器训练时分类器要学习背景信息，训练阶段加入余弦窗后，在目标框中能学习的背景信息变得很有限，进一步降低了滤波器的判别能力。为了抑制边界效应，提升滤波器的性能，研究者提出了一些改进相关滤波跟踪算法。例如，背景感知相关滤波（background aware correlation filter，BACF）算法（Galoogahi等，2017）使用扩大搜索范围裁剪小样本的方法增加真实负样本的数量，在一定程度上缓解了边界效应。背景与方向感知相关滤波（姜文涛等，2021）在BACF基础上利用卡尔曼滤波预测目标运动方向，并提取运动方向上的背景信息，增加滤波器的判别能力。畸变抑制相关滤波（aberrance repressed correlation filters，ARCF）算法（Huang等，2019）在BACF的基础之上通过添加限制响应图变化率的正则化项，在缓解边界效应的同时有效抑制畸变。时间正则化背景感知相关滤波（temporal regularized correlation filters，TRBACF）算法（Yuan等，2020）在BACF算法的基础之上引入时间正则化项，提高算法适应复杂跟踪场景的能力；空间正则化相关滤波跟踪（spatially regularized discriminative correlation filter，SRDCF）算法（Danelljan等，2015）引入空间正则化来解决边界效应问题，有效地抑制背景区域的响应；自适应空间正则化相关滤波（adaptive spatial regularized correlation filter，ASRCF）算法（Dai等，2019）使用自适应空间正则化，进一步缓解了边界效应，在复杂场景下获得了更好的性能；自适应时空正则化相关滤波算法（automatic spatio-temporal regularization tracker，AutoTrack）（Li等，2020b）利用响应图的局部和全局信息实现自适应时空正则化，在速度和精度上得到进一步的提升；畸变抑制时空相关滤波算法（aberrance suppressed spatio-temporal correlation filters，ASTCF）（Elayaperumal和Joo，2021）为抑制边界效应在ARCF算法基础之上引入时间正则化项；长短时空注意相关滤波（spatial and long-short temporal attention correlation filters，STTCF）算法（Zhao等，2022）为抑制边界效应在传统空间正则化项中加入权值矩阵，并且分别设计了两个长短时间正则化项进一步提高跟踪器性能。基于深度学习正则化的相关滤波跟踪算法（蒲磊 等，2020）通过对损失函数的改进来缓解边界效应，有效应对多种复杂场景。而基于注意力学习的正则化相关滤波跟踪算法（仇祝令 等，2020）将掩膜视为注意力权重与滤波器联合求解，能够有效应对边界效应，实现实时准确的跟踪。

上述改进算法虽然在一定程度上抑制了边界效应问题，但是在训练过程中仍然采用了余弦窗操作。而余弦窗会导致提取目标物体的范围缩小，获取的信息有限，大量过滤掉原本需要学习的背景信息，甚至会引入周围噪声导致样本污染（Li等，2020a）。

为解决该问题，1）本文在正则化相关滤波框架中嵌入高斯形状掩膜，重新定义中心和边界的样本权重，在有效降低边界效应的同时，降低余弦窗可能带来的样本污染；2）使用交替方向乘子算法迭代优化算法模型，求得滤波器和空间权重的闭合解；3）在多个基准数据集上进行了大量的对比实验，结果表明嵌入高斯形状掩膜的相关滤波跟踪算法表现出更好的跟踪性能，能有效应对各种具有挑战的场景。本文算法源代码发布于https://github.com/doubleYB2/GSASRCF。

1　相关基线工作

经典的相关滤波跟踪算法使用目标区域循环移位来增加训练样本数目，以此来提高跟踪性能。但是目标区域作为正样本不能范围过大，范围过大会引入不必要的背景信息，导致正样本中无法准确获取有用的目标物体信息，降低跟踪性能。

引入空间正则化是解决边界效应的有效手段。SRDCF 跟踪算法（Danelljan 等，2015）在相关滤波器中应用正则化的方式加入了惩罚项，同时使用不同尺寸搜索解决了尺度变化的问题，在大范围运动等复杂场景中获得较好的性能。算法的目标函数为

$$E(\boldsymbol{w}) = \frac{1}{2}\left\| \sum_{k=1}^{K} \boldsymbol{x}_k * \boldsymbol{w}_k - \boldsymbol{y} \right\|_2^2 + \frac{\lambda}{2}\sum_{k=1}^{K}\left\| \tilde{\boldsymbol{h}} \odot \boldsymbol{w}_k \right\|_2^2 \tag{1}$$

式中，$*$ 为卷积运算，$\odot$ 为逐元素点乘运算，$\boldsymbol{x}_k \in \mathbf{R}^T$、$\boldsymbol{w}_k \in \mathbf{R}^D$ 分别为训练样本和滤波器的第 k 个通道，K 为特征通道数，T 表示采样空间维度，D 表示滤波器维度。$\tilde{\boldsymbol{h}}$ 为滤波器的正则化系数矩阵。$\boldsymbol{y} \in \mathbf{R}^T$ 表示期望的相关响应，λ 为正则化参数。

为了有效利用背景信息训练滤波器，BACF 算法（Galoogahi 等，2017）提取了更大的搜索区域，将整个大的搜索区域进行循环移位之后，用二值掩码矩阵对非目标区域进行裁剪操作。通过该操作得到了大量的真实的负样本，在扩大搜索区域的同时提高了样本的质量和数量，正样本中包含目标物体和周围少数环境信息，而负样本中包含了较多的背景信息。BACF 算法的目标函数为

$$E(\boldsymbol{w}) = \frac{1}{2}\sum_{j=1}^{D}\left\| \sum_{k=1}^{K} \boldsymbol{w}_k^{\mathrm{T}} \boldsymbol{P}\boldsymbol{x}_k[\Delta\tau_j] - \boldsymbol{y}(j) \right\|_2^2 + \frac{\lambda}{2}\sum_{k=1}^{K}\left\| \boldsymbol{w}_k \right\|_2^2 \tag{2}$$

式中，$\boldsymbol{P}$ 是 $D \times T$ 的二值裁剪矩阵，裁剪 $\boldsymbol{x}_k$ 中的 D 个像素。$\boldsymbol{y}(j)$ 是 $\boldsymbol{y}$ 的第 j 个元素，$[\Delta\tau_j]$ 是循环移位算子，$\boldsymbol{x}_k[\Delta\tau_j]$ 表示对 $\boldsymbol{x}_k$ 采用 j 步循环移位。

在跟踪过程中，BACF 和 SRDCF 对不同的目标物体其正则化项保持不变，不能很好地反映目标物体的特征，因此 ASRCF 算法（Dai 等，2019）引入自适应空间正则化，同时优化滤波器系数和空间正则化权值，可以更有效地学习目标物体的特征和更新特定的对象，在跟踪过程中获得可靠的滤波系数。ASRCF 同时利用了两个相关滤波模型进行跟踪，使用一个复杂的相关滤波器对目标物体进行定位的同时，再使用另一个具有多尺度手工特征的相关滤波器进行尺度估计，跟踪效果明显提升，优于其他先进的跟踪算法。自适应空间正则化相关滤波算法的目标计算为

$$E(\boldsymbol{W}, \boldsymbol{h}) = \frac{1}{2}\left\| \sum_{k=1}^{K} \boldsymbol{x}_k * (\boldsymbol{P}^{\mathrm{T}}\boldsymbol{w}_k) - \boldsymbol{y} \right\|_2^2 + \frac{\lambda_1}{2}\sum_{k=1}^{K}\left\| \boldsymbol{h} \odot \boldsymbol{w}_k \right\|_2^2 + \frac{\lambda_2}{2}\left\| \boldsymbol{h} - \boldsymbol{h}^q \right\|_2^2 \tag{3}$$

式中，等号右侧第 1 项为岭回归项，$\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k]$ 为训练数据，$\boldsymbol{W} = [\boldsymbol{w}_1, \boldsymbol{w}_2, \cdots, \boldsymbol{w}_k]$ 为滤波器。第 2 项在滤波器上引入自适应空间正则化，$\boldsymbol{h}$ 为空间权重。第 3 项为了防止模型退化引入了 $\boldsymbol{h}$ 的先验信息 $\boldsymbol{h}^q$ 作为参考权值。λ_1 和 λ_2 分别为第 2 项和第 3 项的正则化参数。将目标公式转换至频域，得到

$$\begin{aligned} E(\boldsymbol{W}, \hat{\boldsymbol{G}}, \boldsymbol{h}) = {} & \frac{1}{2}\left\| \sum_{k=1}^{K} \hat{\boldsymbol{x}}_k \odot \hat{\boldsymbol{g}}_k - \hat{\boldsymbol{y}} \right\|_2^2 + \\ & \frac{\lambda_1}{2}\sum_{k=1}^{K}\left\| \boldsymbol{h} \odot \boldsymbol{w}_k \right\|_2^2 + \\ & \frac{\lambda_2}{2}\sum_{k=1}^{K}\left\| \boldsymbol{h} - \boldsymbol{h}^q \right\|_2^2 \\ \text{s.t. } & \hat{\boldsymbol{g}}_k = \sqrt{T}\,\boldsymbol{F}\boldsymbol{P}^{\mathrm{T}} w_k,\ k = 1, \cdots, K \end{aligned} \tag{4}$$

式中，$\hat{\boldsymbol{G}}$ 为辅助变量矩阵，$\hat{\boldsymbol{G}} = \left[\hat{g}_1, \hat{g}_2, \cdots, \hat{g}_K\right]$。$\boldsymbol{F}$ 是复基向量的正交 $T \times T$ 矩阵，可以将任何 T 维向量化信号映射到傅里叶域。

2　本文算法

2.1　算法模型

相关滤波算法经常在一个样本不平衡的情况下进行学习，能学习到的负样本有限。余弦窗的使用虽然很好地应对了边界效应，但是进一步限制了学习样本的区域，甚至引入不必要的信息造成样本污染。同时，中心样本的重要性要大于周围样本，目标中心相对于边缘应获取更高的响应，而余弦窗简单地使边缘样本为零。因此，本文引入高斯形状掩膜来重新定义目标区域的权重。高斯形状掩膜定义为

$$S(m, n) = \begin{cases} \mathrm{e}^{-\left(\frac{m}{a\beta}\right)^2 - \left(\frac{n}{b\beta}\right)^2} & |m| \leqslant \frac{A}{2} - \frac{a}{2}, |n| \leqslant \frac{B}{2} - \frac{b}{2} \\ 0 & \text{其他} \end{cases} \tag{5}$$

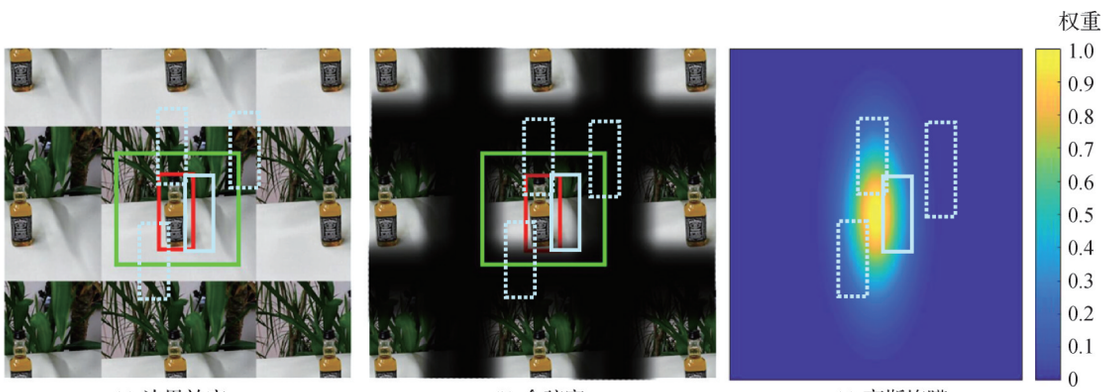
式中，$A \times B$ 是基础图像的大小，$a \times b$ 是目标边界框

的大小，(m,n)是实际图像样本中心的位置，β是控制训练样本权值衰减速度的参数。

在自适应空间正则化相关滤波算法中嵌入高斯形状掩膜（Gaussian shape mask adaptive spatial regularization correlation filter，GSASRCF），对边界效应进一步抑制，在训练时重新对样本的权重进行更新。本文算法目标计算为

$$E(\boldsymbol{W},\hat{\boldsymbol{G}},\boldsymbol{h})=\frac{1}{2}\left\|\boldsymbol{S}\odot\left(\sum_{k=1}^{K}\hat{\boldsymbol{x}}_k\odot\hat{\boldsymbol{g}}_k-\hat{\boldsymbol{y}}\right)\right\|_2^2+\frac{\lambda_1}{2}\sum_{k=1}^{K}\left\|\boldsymbol{h}\odot\boldsymbol{w}_k\right\|_2^2+\frac{\lambda_2}{2}\left\|\boldsymbol{h}-\boldsymbol{h}^q\right\|_2^2$$
$$\text{s.t. }\hat{\boldsymbol{g}}_k=\sqrt{T}\,\boldsymbol{F}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{w}_k,\ k=1,\cdots,K \tag{6}$$

将高斯形状掩膜引入自适应空间正则化相关滤波算法，避免边界不连续，降低余弦窗带来的样本污染问题。如图1所示，图1(a)绿色框中的区域为基础图像区域（即搜索区域），大小为$A\times B$，红色框区域为目标区域，大小为$a\times b$，在基础图像区域获取训练样本时，收集的样本可能会超出基础图像区域，引起边界不连续问题（即边界效应），如图中青色虚线框所示的样本，覆盖了边界部分，出现边界不连续问题。图1(b)中余弦窗和目标区域相乘使边缘像素为零，压缩了提取样本的区域，使得训练样本受到周围噪声的污染（样本中的黑色区域），且靠近中心的样本和边缘的样本的权重均相同。图1(c)中高斯形状掩膜抑制边界负样本的同时，对靠近目标中心和基础图像区域边缘的样本的权重重新分配，增加了中心样本的重要性，同时降低远离中心的边缘样本的重要性，增加中心样本的响应，从而降低样本污染的影响，增强滤波器的判别能力。



(a) 边界效应　(b) 余弦窗　(c) 高斯掩膜

图1　高斯形状掩膜重新分配权重可视化

Fig. 1　Visualization of the Gaussian shape mask for redistributing the sample weight

((a)boundary effect;(b)cosine window;(c)Gaussian mask)

2.2　模型求解

对本文算法的目标公式使用增广拉格朗日方法进行求解，式(6)利用增广拉格朗日法，可得

$$L\left(\boldsymbol{W},\hat{\boldsymbol{G}},\boldsymbol{h},\hat{\boldsymbol{V}}\right)=D\left(\boldsymbol{W},\hat{\boldsymbol{G}},\boldsymbol{h}\right)+\sum_{k=1}^{K}\hat{\boldsymbol{v}}_K^{\mathrm{T}}\left(\hat{\boldsymbol{g}}_k-\sqrt{T}\,\boldsymbol{F}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{w}_k\right)+\frac{\mu}{2}\sum_{k=1}^{K}\left\|\hat{\boldsymbol{g}}_k-\sqrt{T}\,\boldsymbol{F}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{w}_k\right\|_2^2 \tag{7}$$

式中，$\boldsymbol{V}=[\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_k]\in\mathbf{R}^{T\times K}$是拉格朗日乘子，$\mu$为交替方向乘子法（alternating direction of multiplier method，ADMM）惩罚因子。

通过引入变量$\boldsymbol{u}_k=\frac{1}{\mu}\boldsymbol{v}_k$，$k=1,2,\cdots,K$，式(7)可转换为

$$L\left(\boldsymbol{W},\hat{\boldsymbol{G}},\boldsymbol{h},\hat{\boldsymbol{U}}\right)=\frac{1}{2}\left\|\boldsymbol{S}\odot\left(\sum_{k=1}^{K}\hat{\boldsymbol{x}}_k\odot\hat{\boldsymbol{g}}_k-\hat{\boldsymbol{y}}\right)\right\|_2^2+\frac{\lambda_1}{2}\sum_{k=1}^{K}\left\|\boldsymbol{h}\odot\boldsymbol{w}_k\right\|_2^2+\frac{\lambda_2}{2}\sum_{k=1}^{K}\left\|\boldsymbol{h}-\boldsymbol{h}^q\right\|_2^2+\frac{\mu}{2}\sum_{k=1}^{K}\left\|\hat{\boldsymbol{g}}_k-\sqrt{T}\,\boldsymbol{F}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{w}_k+\hat{\boldsymbol{u}}_k\right\|_2^2 \tag{8}$$

式中，$\hat{\boldsymbol{U}}=\left[\hat{\boldsymbol{u}}_1,\hat{\boldsymbol{u}}_2,\cdots,\hat{\boldsymbol{u}}_K\right]\in\mathbf{R}^{T\times K}$。使用ADMM对算法进行优化，将原问题转化为关于滤波器$\boldsymbol{W}$和辅助变量$\boldsymbol{G}$的两个子问题，并且所有的子问题都有封闭解。

1）求解子问题$\boldsymbol{W}$。$\boldsymbol{W}$的最优解为

$$\boldsymbol{w}_k^*=\arg\min_{\boldsymbol{w}_k}\left\{\frac{\lambda_1}{2}\left\|\boldsymbol{h}\odot\boldsymbol{w}_k\right\|_2^2+\frac{\mu}{2}\left\|\hat{\boldsymbol{g}}_k-\sqrt{T}\,\boldsymbol{F}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{w}_k+\hat{\boldsymbol{u}}_k\right\|_2^2\right\}=[\lambda_1\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}+\mu T\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}]^{-1}\mu T\boldsymbol{P}(\boldsymbol{u}_k+\boldsymbol{g}_k) \tag{9}$$

式中，$\boldsymbol{H} = \mathrm{diag}(\boldsymbol{h})$表示一个对角矩阵，将$\boldsymbol{h}$中所有元素进行对角化。

2）求解子问题$\hat{\boldsymbol{G}}^*$。$\boldsymbol{G}$的最优解为

$$\hat{\boldsymbol{G}}^* = \arg\min_{\hat{\boldsymbol{G}}} \left\{ \frac{1}{2} \left\| \boldsymbol{S} \odot \left(\sum_{k=1}^{K} \hat{\boldsymbol{x}}_k \odot \hat{\boldsymbol{g}}_k - \hat{\boldsymbol{y}} \right) \right\|_2^2 + \frac{\mu}{2} \sum_{k=1}^{K} \left\| \hat{\boldsymbol{g}}_k - \sqrt{T}\, \boldsymbol{F}\boldsymbol{P}^{\mathrm{T}} \boldsymbol{w}_k + \hat{\boldsymbol{u}}_k \right\|_2^2 \right\} \tag{10}$$

由于该式计算比较复杂，将问题进行拆分，对每个像素的所有通道进行处理，对问题进行优化，具体为

$$\boldsymbol{v}_j^*(\hat{\boldsymbol{G}}) = \arg\min_{\boldsymbol{v}_j(\hat{\boldsymbol{G}})} \left\{ \frac{1}{2} \left\| \boldsymbol{S} \odot (\boldsymbol{v}_j^{\mathrm{T}}(\hat{\boldsymbol{X}}) \boldsymbol{v}_j(\hat{\boldsymbol{G}}) - \hat{\boldsymbol{y}}_j) \right\|_2^2 + \frac{\mu}{2} \sum_{k=1}^{K} \left\| \boldsymbol{v}_j(\hat{\boldsymbol{G}}) + \boldsymbol{v}_j(\hat{\boldsymbol{M}}) \right\|_2^2 \right\} \tag{11}$$

式中，$\boldsymbol{v}_j(\hat{\boldsymbol{M}}) = \boldsymbol{v}_j(\hat{\boldsymbol{U}}) - \boldsymbol{v}_j(\sqrt{T}\, \boldsymbol{F}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{W})$。由于求得的解中带有求逆运算，求解的计算量依旧庞大，因此使用 Sherman-Morrison 公式 $(\boldsymbol{A} + \boldsymbol{u}\boldsymbol{v}^{\mathrm{T}})^{-1} = \boldsymbol{A}^{-1} - \frac{\boldsymbol{A}^{-1}\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}\boldsymbol{A}^{-1}}{1 + \boldsymbol{v}^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{u}}$对计算进行加速，求得最后的闭合解为

$$\boldsymbol{v}_j^*(\hat{\boldsymbol{G}}) = \frac{1}{\mu T} \left(\boldsymbol{I} - \frac{\boldsymbol{v}_j(\hat{\boldsymbol{X}}) \boldsymbol{S}^{\mathrm{T}} \boldsymbol{S} \boldsymbol{v}_j^{\mathrm{T}}(\hat{\boldsymbol{X}})}{\mu T + \boldsymbol{v}_j^{\mathrm{T}}(\hat{\boldsymbol{X}}) \boldsymbol{S}\boldsymbol{S}^{\mathrm{T}} \boldsymbol{v}_j(\hat{\boldsymbol{X}})} \right) \times \left(\hat{y}_j \boldsymbol{S}\boldsymbol{S}^{\mathrm{T}} \boldsymbol{v}_j(\hat{\boldsymbol{X}}) + \mu \boldsymbol{v}_j(\sqrt{T}\, \boldsymbol{F}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{W}) - \mu \boldsymbol{v}_j(\hat{\boldsymbol{U}}) \right) \tag{12}$$

得到上述两个子问题的解之后，$\boldsymbol{W}$、$\hat{\boldsymbol{G}}$和$\hat{\boldsymbol{U}}$固定，可求得权重$\boldsymbol{h}$的闭合解为

$$\boldsymbol{h}^* = \arg\min_{\boldsymbol{h}} \left\{ \frac{\lambda_1}{2} \sum_{k=1}^{K} \left\| \boldsymbol{R}_k \boldsymbol{h} \right\|_2^2 + \frac{\lambda_2}{2} \left\| \boldsymbol{h} - \boldsymbol{h}^b \right\|_2^2 \right\} = \left(\lambda_1 \sum_{k=1}^{K} \boldsymbol{R}_k^{\mathrm{T}} \boldsymbol{R}_k + \lambda_2 \boldsymbol{I} \right)^{-1} \lambda_2 \boldsymbol{h}^b \tag{13}$$

式中，$\boldsymbol{R}_k = \mathrm{diag}(\boldsymbol{w}_k)$。拉格朗日乘子更新为

$$\hat{\boldsymbol{U}}^{i+1} = \hat{\boldsymbol{U}}^i + \hat{\boldsymbol{G}}^{i+1} - \hat{\boldsymbol{W}}^{i+1} \tag{14}$$

式中，$\hat{\boldsymbol{U}}^i$为第i次迭代的拉格朗日乘子的傅里叶变换，$\hat{\boldsymbol{G}}^{i+1}$和$\hat{\boldsymbol{W}}^{i+1}$是上面两个子问题的迭代$i+1$次的当前解。

2.3　目标位置与尺度估计

将相关滤波器的响应定义为$\boldsymbol{z}$，在频域中确定$\boldsymbol{z}$的最大值即为目标的估计位置，即

$$\hat{\boldsymbol{z}}_m = \sum_{k=1}^{K} \hat{\boldsymbol{x}}_k \odot \hat{\boldsymbol{g}}_k \tag{15}$$

本文使用手工特征与深度特征的集成特征来训练目标位置估计的相关滤波器（位置估计 CF）。其中，使用的手工特征是方向梯度直方图（histogram of gradient，HOG）特征，深度特征来自 VGG-M 网络 Norm1 层和 VGG-16 的 Conv4-3 层，跟踪过程仅在一个尺度搜索区域提取这些特征。

与经典的相关滤波跟踪算法一样，模板更新方法为

$$\hat{\boldsymbol{X}}_m^{\mathrm{new}} = (1 - \eta)\hat{\boldsymbol{X}}_m^{\mathrm{old}} + \eta \hat{\boldsymbol{X}}^* \tag{16}$$

式中，$\hat{\boldsymbol{X}}_m^{\mathrm{old}}$是旧的模板模型，$\hat{\boldsymbol{X}}_m^{\mathrm{new}}$是更新的模板模型，$\hat{\boldsymbol{X}}^*$是当前帧的模板，$\eta$是在线学习率。

对于目标的尺度估计，本文单独训练一个尺度估计相关滤波器（尺度 CF）。尺度估计滤波器使用手工 HOG 特征进行训练，跟踪过程中使用 5 个尺度进行尺度估计，根据 5 个响应图的最高分数确定最佳的目标尺度。GSASRCF 算法整体流程如图 2 所示，其中实线箭头代表训练过程，虚线箭头代表跟踪过程，FFT（fast Fourier transform）表示快速傅里叶变换，IFFT（inverse FFT）表示逆快速傅里叶变换。算法的步骤如下：

算法 1：GSASRCF 算法

输入：视频序列，第 1 帧初始目标位置及尺度（宽度、高度）。

输出：序列每帧中目标的位置及尺度。

1）初始化：根据初始目标位置及尺度，确定搜索区域，计算高斯掩膜、正则化权重、定位 CF 和尺度 CF，确定初始滤波器模板。

2）对于序列的第 2 帧至最后一帧：

（1）根据上一帧目标位置及尺度估计结果，确定搜索区域。

（2）提取不同尺度的 HOG 特征及深度特征，构建目标描述。

（3）将上一帧的定位 CF 和尺度 CF 与当前帧的目标描述做相关运算，求得最大响应，确定目标位置及尺度，输出估计目标位置及尺度。

（4）根据当前帧的跟踪结果，确定搜索区域。

（5）提取训练样本，并提取 5 个尺度的 HOG 特征，使用高斯掩膜参与 ADMM 训练尺度 CF。

（6）提取训练样本，并提取单尺度 HOG 特征及深度特征，使用高斯掩膜参与 ADMM 方法训练定位 CF。

（7）根据学习率更新滤波器模板。

3）结束。

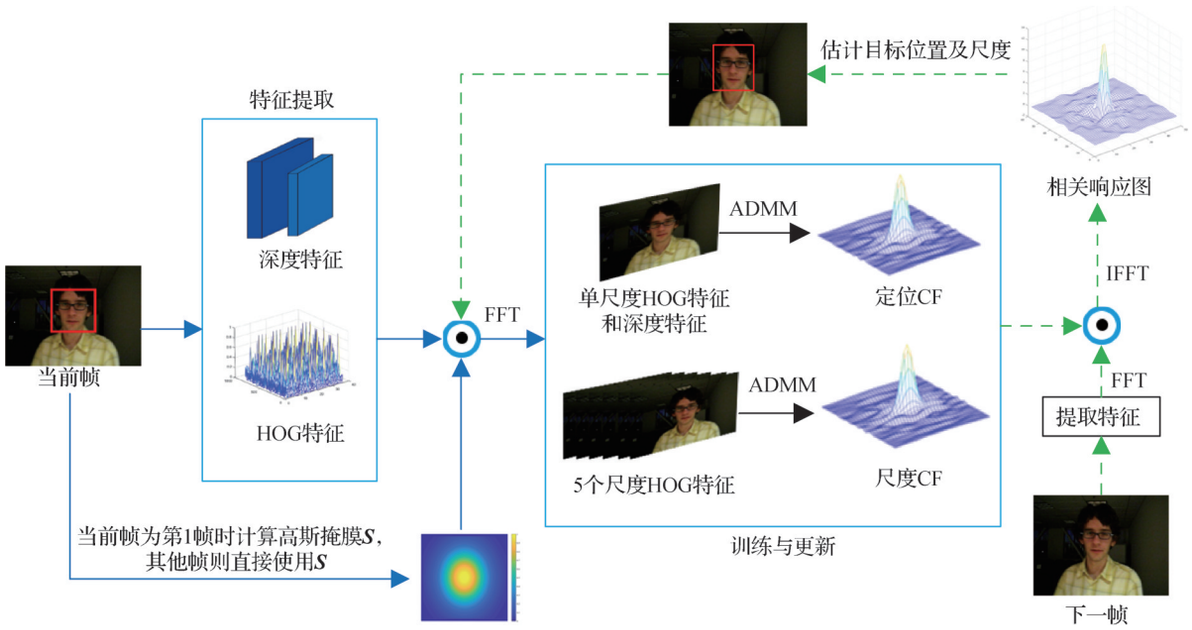


图2　GSASRCF算法图示

Fig. 2　Illustration of GSASRCF

3　实验结果与分析

3.1　实验细节

本文使用MATLAB 2018a实现GSASRCF算法，所有实验均在64位Windows10平台下完成。计算机处理器参数为Intel(R) Core (TM) i7-9700 3.00 GHz CPU，内存为48 GB，图形显卡为NVIDIA Quadro P400。正则化参数$\lambda_1 = 1.2$，$\lambda_2 = 0.001$。使用VGG（Visual Geometry Group）网络特征（VGG-M的Norm1，VGG-16的Conv4-3）和HOG特征进行目标定位，使用5个尺度的HOG特征进行尺度估计。控制训练样本权值衰减速度的参数$\beta = 4$，在线学习率$\eta = 0.0185$。ADMM迭代次数设置为2，惩罚因子$\mu = 1$。

本文使用OTB2013（online object tracking benchmark）（Wu等，2013）、TC128（temple color）（Liang等，2015）、UAV123（unmanned aerial vehicle）（Mueller等，2016）和Got-10k（general object tracking 10 000）（Huang等，2021）4个公开基准数据集进行跟踪器评测。除Got-10k数据集外，将本文提出的GSASRCF算法与其他相关滤波跟踪算法采用一遍评估法（one pass evaluation，OPE）进行评估，使用精确度和成功率对算法进行测评。使用中心误差作为精确度的评价指标，即跟踪目标物体的中心位置与手动标记的目标物体真实值之间的平均欧氏距离。使用序列所有帧的中心误差的平均值作为标准，评价本文算法在整个序列上的跟踪性能。定义阈值为20像素，预测值与真实值差距在阈值范围内的帧数占总帧数的百分比可以绘制出一条曲线，从而生成精度曲线。跟踪预测的边界框定义为$\boldsymbol{r}_t$，真实标定的边界框定义为$\boldsymbol{r}_a$，重叠率定义为$\boldsymbol{O} = \frac{\boldsymbol{r}_t \cap \boldsymbol{r}_a}{\boldsymbol{r}_t \cup \boldsymbol{r}_a}$，这里的∩和∪分别代表这两个区域的交集和并集。当$\boldsymbol{O}$大于给定阈值时即视为成功，计算跟踪成功的帧数占总帧数的比率即为成功率。但是在一个特定的阈值上计算成功率不具有代表性，采用曲线下面积（area under curve，AUC）来对跟踪算法进行排序。

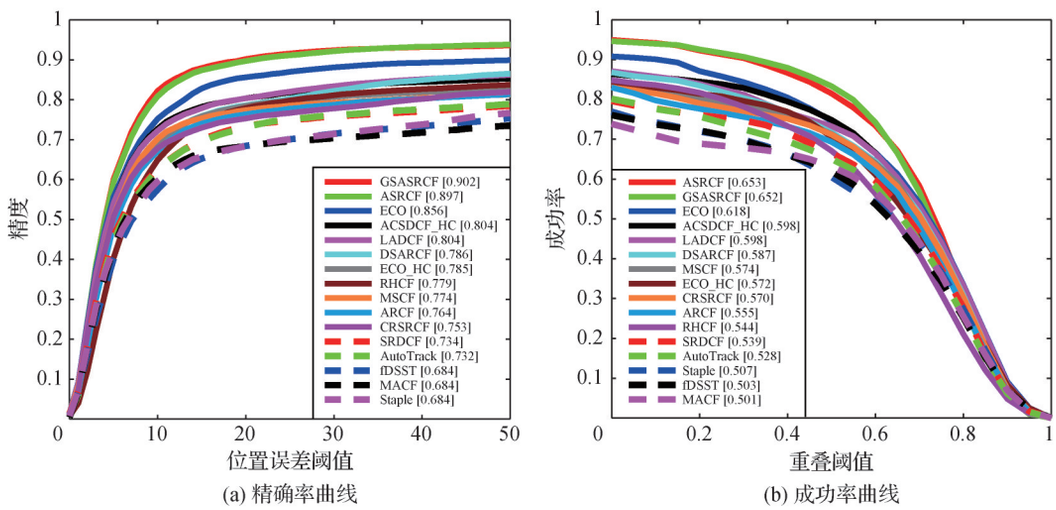
3.2　定量分析

3.2.1　OTB2013数据集

OTB2013数据集中包含50个视频序列，其中包含部分灰度视频序列。该数据集包含了11种挑战因素，包括快速运动（fast motion，FM）、背景杂波（background clutter，BC）、运动模糊（motion blur，MB）、形变（deformation，DEF）、光照变化（illumination variation，IV）、平面内旋转（in-plane rotation，IPR）、低分辨率（low resolution，LR）、尺度变化（scale variation，SV）、平面外旋转（out-of-plane rotation，OPR）、遮挡（occlusion，OCC）和移出视野（out of

view，OV）。对比算法包括 ARCF（Huang 等，2019）、SRDCF（Danelljan 等，2015）、AutoTrack（Li 等，2020b）、ECO（efficient convolution operators）（Danelljan 等，2017a）、ACSDCF_HC（adaptive channel selection discriminative CF with hand-crafted feature）（Xu 等，2021）、RHCF（robust hierachical CF）（曾梦媛 等，2020）、MSCF（mutation sensitive CF）（Zheng 等，2021）、MACF（motion-aware CF）（Zhang 等，2018）、ECO_HC（ECO with hand-crafted feature）（Danelljan 等，2017a）、DSARCF（dynamic saliency aware regularized CF）（Feng 等，2019）、LADCF（learning adaptive discriminative CF）（Xu 等，2019）、CRSRCF（content-related spatial regularization CF）（Han 等，2018）、fDSST（fast discriminative scale space tracking）（Danelljan 等，2017b）、Staple（sum of template and pixel-wise learners）（Bertinetto 等，2016）和 ASRCF（adaptive spatial regularized CF）（Dai 等，2019）。

图 3 展示了本文算法与其他对比相关滤波跟踪算法的 OPE 结果。在精度曲线图中，本文算法 GSASRCF 排在第 1 位。与基准算法 ASRCF 相比提高 0.5%。GSASRCF 精度得分 0.902，超过第 3 名 ECO 算法（0.856）4.6%。ACSDCF_HC（0.804）和 LADCF（0.804）以相同分数排在第 4 位，使用显著性特征的 DSASRCF（0.786）排在精度算法对比的第 6 位。在成功率得分中，GSASRCF 算法 AUC 得分 0.652，排名第 2，比 ASRCF 算法低 0.1%，相比 ECO 算法（0.618）获得了 3.4% 的增益，同时相比排名第 3 的 ACSDCF_HC（0.598）和 LADCF（0.598）高出了 5.4%。相比 MSCF（0.574）、ARCF（0.555）、RHCF（0.544）等也取得了极具竞争力的结果。本文算法 GSASRCF 达到了最好的精度，引入高斯形状掩膜的算法在算法精度对比中展现了良好的跟踪性能。



(a) 精确率曲线　　(b) 成功率曲线

图 3　OTB2013 数据集的对比结果

Fig. 3　Comparison results on OTB2013 dataset ((a) precision plots; (b) success plots)

表 1 展示了本文算法 GSASRCF 与其他相关滤波算法在 11 个不同挑战场景中的 AUC 得分结果。在快速运动（0.644）、背景杂波（0.659）、形变（0.632）、尺度变化（0.632）、平面外旋转（0.646）、遮挡（0.640）、超出视野（0.649）等 7 个场景中，本文算法 AUC 得分排名第 1。其他 4 个挑战场景均排在第 2 位。使用深度特征的 ECO 算法在 9 个挑战场景中排名第 3 位，在运动模糊场景（0.648）取得高分数。GSASRCF 算法在 11 种不同挑战场景中展现了其优越性，加入高斯形状掩膜的相关滤波器可以有效地应对边界效应，降低引入余弦窗所导致的样本污染的影响，更好地处理目标形变、遮挡和尺度变化等挑战，提高跟踪器的跟踪性能。

3.2.2　TC128 数据集

TC128 数据集由 128 个彩色序列组成，包含的 11 个挑战因素与 OTB 数据集相同。本文算法与其他先进的目标跟踪算法进行 OPE 测评，包括 ARCF、ACSDCF_HC、MSCF（Zheng 等，2021）、ECO_HC、STRCF（Li 等，2018）、MRCF（Ye 等，2022）、AutoTrack（Li 等，2020b）、RHCF、DSARCF、Staple、DRCF（Fu 等，2020）、SRDCF、DSST（Danelljan 等，2017b）和 ASRCF。

表1　OTB2013数据集中11个不同挑战场景的对比结果(AUC)

Table 1　Comparisons of the 11 challenging attributes on OTB2013(AUC)

跟踪算法	FM	BC	MB	DEF	IV	IPR	LR	SV	OPR	OCC	OV
ASRCF (Dai等,2019)	0.641	0.657	0.633	0.628	**0.665**	**0.627**	**0.595**	0.629	0.641	0.632	0.637
ECO (Danelljan等,2017a)	0.639	0.592	**0.648**	0.554	0.612	0.574	0.567	0.617	0.588	0.621	0.599
LADCF (Xu等,2019)	0.576	0.576	0.572	0.529	0.551	0.545	0.546	0.590	0.561	0.583	0.560
DSARCF (Feng等,2019)	0.572	0.602	0.568	0.525	0.583	0.547	0.549	0.586	0.557	0.552	0.506
ECO_HC (Danelljan等,2017a)	0.576	0.562	0.535	0.510	0.523	0.506	0.566	0.557	0.529	0.549	0.514
ACSDCF_HC (Xu等,2021)	0.596	0.604	0.595	0.517	0.577	0.559	0.545	0.598	0.573	0.596	0.545
RHCF (曾梦媛 等,2020)	0.549	0.515	0.556	0.499	0.542	0.540	0.513	0.521	0.512	0.509	0.460
MSCF (Zheng等,2021)	0.565	0.562	0.567	0.549	0.561	0.553	0.496	0.533	0.541	0.521	0.530
AutoTrack (Li等,2020b)	0.536	0.505	0.534	0.514	0.525	0.479	0.529	0.486	0.493	0.506	0.521
MACF (Zhang等,2018)	0.496	0.555	0.510	0.418	0.565	0.468	0.477	0.477	0.445	0.444	0.456
ARCF (Huang等,2019)	0.556	0.535	0.545	0.548	0.536	0.505	0.512	0.508	0.504	0.518	0.488
CRSRCF (Han等,2018)	0.595	0.568	0.583	0.496	0.550	0.519	0.501	0.564	0.519	0.534	0.484
SRDCF (Danelljan等,2015)	0.562	0.530	0.539	0.466	0.534	0.486	0.495	0.519	0.485	0.498	0.431
fDSST (Danelljan等,2017b)	0.532	0.543	0.542	0.434	0.549	0.490	0.458	0.483	0.453	0.446	0.454
Staple (Bertinetto等,2016)	0.483	0.476	0.476	0.511	0.516	0.462	0.399	0.460	0.463	0.489	0.455
GSASRCF (本文)	**0.644**	**0.659**	0.635	**0.632**	0.662	0.625	0.593	**0.632**	**0.646**	**0.640**	**0.649**

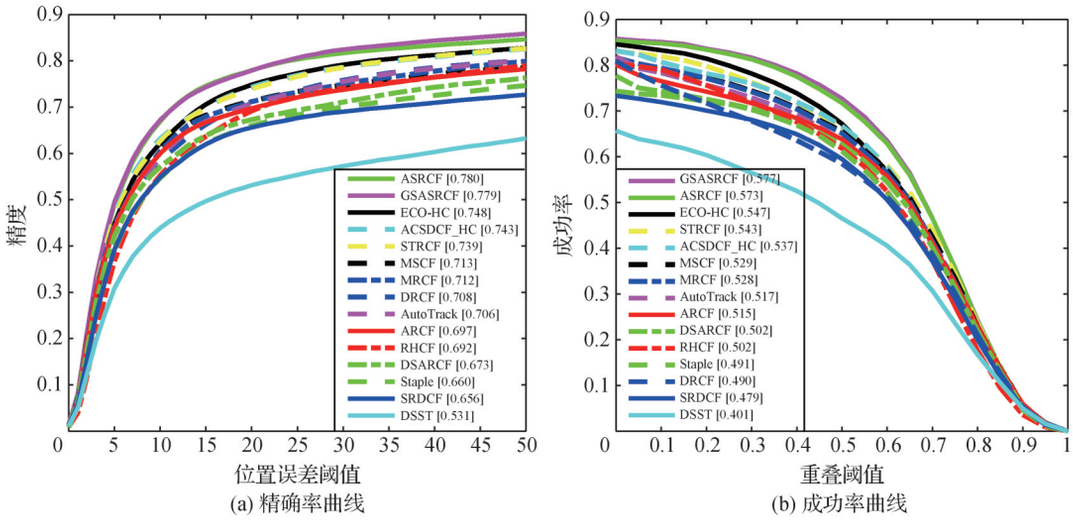
注:加粗字体表示各列最优结果。

图4展示了在TC128数据集中各跟踪算法的精确度排名和根据AUC得分生成的成功率排名。GSASRCF算法在平均精度得分中排名第2,排名第1的是ASRCF(0.780),本文算法与ASRCF算法差距仅为0.001。相比其他跟踪算法取得了比较好的结果。本文算法在总的成功率图中得分0.577,排在对比算法的第1名。基准算法ASRCF(0.573)排在第2位,比本文算法低0.4%。ECO-HC算法AUC得分0.547和STRCF算法AUC得分0.543,分别排在第3名和第4名,分别比GSASRCF算法低3%和3.4%。综合精度和成功率排名,嵌入高斯形状掩膜的算法可以更好抵抗物体的形变等挑战因素,本文算法在TC128数据集中也表现出很强的竞争力。

3.2.3　UAV123数据集

UAV123数据集中包含无人机拍摄的123个彩色序列,与OTB2013和TC128一样采用精确度和重叠率对算法进行排名,但无人机数据集挑战场景与其他两个数据集不相同,包含12种挑战场景:纵横比变换(aspect ratio change, ARC)、背景杂波(BC)、相机运动(camera motion, CM)、快速运动(FM)、全遮挡(full occlusion, FOC)、光照变化(IV)、低分辨率(LR)、移出视野(OV)、部分遮挡(partial occlusion, POC)、相似物体(similar object, SOB)、尺度变化(SV)和视角变化(viewpoint change, VC)。本文算法在UAV123数据集与AutoTrack、ACSDCF_HC、ECO_HC、ADTrack_v2(Li等,2023)、DRCF、MSCF、STRCF、SRDCF、ARCF、Staple、CRSRCF、MACF、MRCF和ASRCF等14个先进的CF跟踪算法进行对比。

图5显示了算法在UAV123无人机数据集上的精度和成功率对比结果。该数据集中算法对比结果与前两个数据集排名有所不同。本文算法GSASRCF得分0.741,在精度对比中排名第1。基础算法ASRCF(0.738)排在第2位,比GSASRCF算法低0.3%。ECO_HC得分0.725,排在第3位,低于GSASRCF算法1.6%。专门针对无人机目标的跟踪算法ADTrack_v2(0.701)、DRCF(0.700)、MRCF(0.693)、MSCF(0.690)和AutoTrack(0.689)分别排在第4～8名。而在OTB2013排名第4位表现相对良好的ACSDCF_HC算法在UAV123数据集中仅排

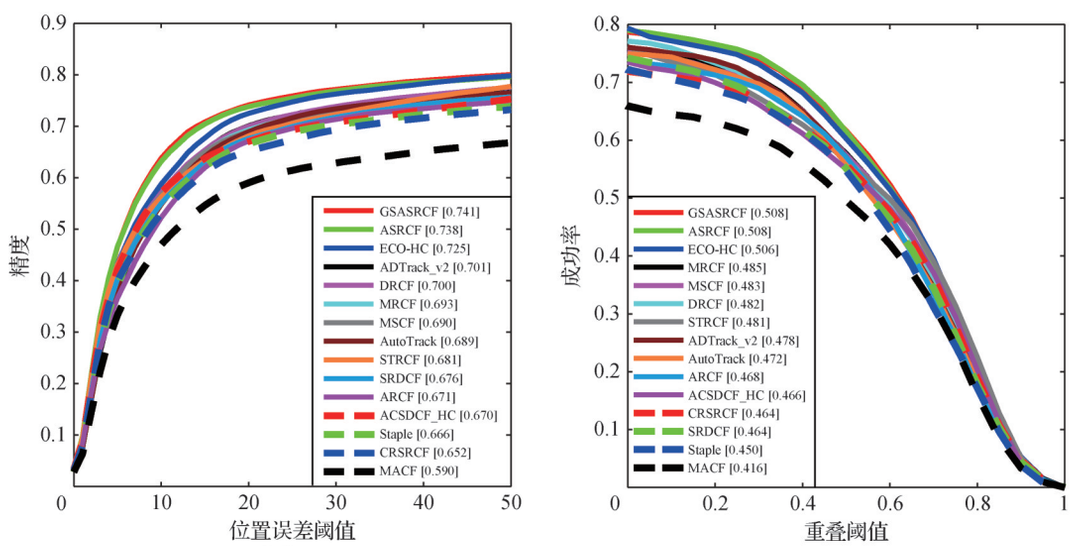


(a) 精确率曲线　　(b) 成功率曲线

图 4　TC128 数据集的对比结果

Fig. 4　Comparison results on TC128 dataset ((a) precision plots; (b) success plots)

名第 12 位，精度得分 0.670，说明该算法难以应对无人机目标跟踪中的挑战。成功率 AUC 得分排名与精度排名略有差距，本文算法 AUC 得分 0.508 排名第 1，与 ASRCF 取得相同的分数。ECO-HC 算法以 0.506 排在第 2 名，MRCF（0.485）和 MSCF（0.483）分列第 3、4 位。精度排名第 4 位的 ADTrack_v2 算法的 AUC 得分为 0.472，排在第 7 位。本文算法在 UAV 数据集中仍表现良好，较前两个数据集没有明显波动，在精度和 AUC 得分中排在 15 个算法的第 1 名，取得了相当具有竞争力的结果，说明引入高斯形状掩膜能够提高算法处理无人机目标跟踪挑战的能力。



(a) 精确率曲线　　(b) 成功率曲线

图 5　UAV123 数据集的对比结果

Fig. 5　Comparison results on UAV123 dataset ((a) precision plots; (b) success plots)

4.2.4　Got-10k 数据集

Got-10k 数据集包含了 84 类移动物体和 32 种运动模式，具有较强的挑战性。该数据集使用平均重叠率（average overlap，AO）和成功率（success rate，SR）指标与其他 14 个算法进行对比测评，其中 AO 表示真实值和估计边界框之间重叠的平均值。SR 为重叠超过阈值的成功跟踪帧的百分比。如图 6 所示，本文算法在 AO 分数上优于其他对比算法，排名第 1。RHCF 算法（0.314）排在对比算法第 2 位。本文所依托的基础算法 ASRCF（0.313）排在第 3 位，

MRCF(0.299)和MSCF(0.292)分列第4、5位。本文算法使用自适应空间正则化算法嵌入高斯形状掩膜，在AO指标上有明显的提升。表2列出了所有对比算法的AO和SR值。在成功率SR方面，阈值为0.5时，GSASRCF算法成功率为0.319，位列第1，ASRCF和RHCF分列2、3位。上述结果充分表明，本文引入高斯形状掩膜，提高了相关滤波跟踪算法处理不同运动模式挑战的能力。

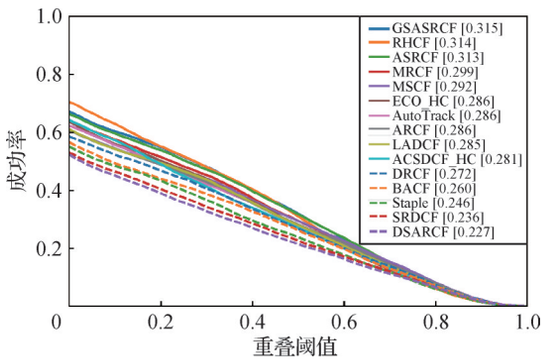


图6　Got-10k数据集的对比结果

Fig. 6　Comparison results on Got-10k dataset

在该实验中同时对算法的帧率(frame per second, FPS)进行了对比，如表2所示，使用手工特征的ECO_HC算法速度最快44.55帧/s，其他采用手工特征的算法，如fDSST(30.43帧/s)和ACSDCF_HC(29.08帧/s)，也具有较高的计算效率。而使用深度网络特征的目标跟踪算法虽然成功率较高，但速度略有不足，AO表现较好的RHCF算法仅为1.80帧/s。本文算法GSARCF的速度为5.2帧/s，比基础算法ASRCF(5.43帧/s)速度下降约0.2帧/s，其原因在于引入高斯掩膜后，增加了部分计算时间。

3.3　消融实验

本文GSASRCF跟踪算法在基准算法自适应空间正则化相关滤波算法(ASRCF)基础上引入高斯形状掩膜，训练位置滤波器和尺度滤波器进行跟踪，本节首先在TC128数据集上对在基准算法的不同滤波器中引入高斯形状掩膜的跟踪结果进行了消融实验，对比基线算法ASRCF、在位置滤波器加入高斯掩膜的算法L-GSASRCF、在尺度滤波器加入高斯掩膜的算法S-GSASRCF以及在两个滤波器同时加入高斯掩膜的算法B-GSASRCF，实验结果如表3所示。

表3中的总体成功率结果显示，对于仅在尺度滤波器或位置滤波器上加入高斯掩膜的改进算法，

表2　Got-10k数据集测评结果

Table 2　Evaluation results on Got-10k dataset

跟踪算法	AO	SR(阈值=0.5)	帧率/(帧/s)
DSST (Damelljan等,2017b)	0.247	0.223	18.25
fDSST (Damelljan等,2017b)	0.206	0.187	30.43
SRDCF (Damelljan等,2015)	0.236	0.227	5.58
Staple (Damelljan等,2016)	0.246	0.239	28.87
ASRCF (Dai等,2019)	0.313	0.317	5.43
ECO_HC (Daneujan等,2017a)	0.286	0.276	44.55
MSCF (Zheng等,2021)	0.292	0.294	20.65
AutoTrack (Li等,2020b)	0.286	0.277	**35.83**
DRCF (Fu等,2020)	0.272	0.267	18.65
ARCF (Huang等,2019)	0.286	0.282	16.33
ACSDCF_HC (Xu等,2021)	0.281	0.270	29.08
DSARCF (Feng等,2019)	0.227	0.215	5.88
LADCF (Xu等,2019)	0.285	0.279	16.33
MRCF (Ye等,2022)	0.299	0.294	14.62
RHCF (曾梦媛 等,2020)	0.314	0.310	1.80
GSASRCF (本文)	**0.315**	**0.319**	5.20

注：加粗字体表示各列最优结果。

成功率AUC与基准算法相比出现了不同幅度的下降。而在两个滤波器中同时引入高斯形状掩膜(B-GSASRCF)取得了更好的结果，其总体成功率AUC为0.577，比基准算法ASRCF高0.4%。表3同时对比了不同挑战因素上的成功率数据，对于OPR、OV、IV、FM、IPR等挑战因素，S-GSASRCF算法的结果都高于基准算法，而同时加入高斯掩膜的B-GSASRCF在这6种挑战因素中获得了最高的AUC结果。其原因在于尺度滤波器和位置滤波器都使用了HOG特征，很难处理遮挡问题，当加入高斯形状掩膜时，可以降低边缘像素的权重，因此在两个滤波器中都加入高斯形状掩膜可以提升算法在TC128的全彩序列中的表现。

图7所示为本文算法GSASRCF与基准算法ASRCF在TC128数据集上的定性比较结果。第1行中的目标为快速运动的小球，结果显示小球运动方向和速度发生剧烈变化时，ASRCF算法出现了漂移，而GSASRCF算法则持续地捕获小球。第2行所示场景中发生剧烈光照变化，目标人脸外观发生剧

表 3　TC128 数据集上的消融实验结果（成功率）

Table 3　Ablation experimental results on TC128 dataset（success）

算法	OCC	OPR	OV	IV	FM	IPR	总成功率
ASRCF	0.535	0.569	0.502	0.581	0.528	0.533	0.573
L-GSASRCF	0.515	0.562	0.477	0.568	0.518	0.525	0.562
S-GSASRCF	0.533	0.570	0.515	0.588	0.533	0.536	0.572
B-GSASRCF	**0.539**	**0.573**	**0.522**	**0.590**	**0.535**	**0.537**	**0.577**

注：加粗字体表示各列最优结果。

烈变化，第 95、96 帧 ASRCF 逐渐丢失目标，GSASRCF 仍然能够跟踪到目标。第 3 行序列中包含目标遮挡的挑战，ASRCF 在第 73 帧开始逐渐漂移，GSASRCF 能够较好地捕获目标运动员。第 4 行为平面外旋转的跟踪挑战，第 488 帧目标发生旋转前，ASRCF 和 GSASRCF 均能精确定位目标，但在第 495 帧目标发生旋转后，ASRCF 开始逐步丢失目标，在第 504 帧目标几乎完全丢失。最后一行目标球移出视野后重新进入视野（第 519 帧），可以看到重新进入视野后本文算法重新捕获目标并稳定跟踪，而基准算法 ASRCF 则跟踪失败。

上述定量和定性分析结果表明，在基准算法 ASRCF 中引入高斯形状掩膜后，有效提高了算法处理诸如遮挡、移出视野、平面外旋转和快速运动等挑战的能力。

图 7　本文算法与基准算法在 TC128 数据集上的定性比较

Fig. 7　Qualitative results of our method and baseline method on TC128 dataset

3.4　定性分析

为直观地展示本文算法与其他算法的跟踪能力的对比，进行了可视化的定性分析。如图 8 所示。第 1 行视频序列为 Bird1 序列，本文算法与其他 5 个表现良好的算法进行对比。Bird1 序列中包含形变、遮挡和超出视野的挑战因素。由图 8 可以看出，

第32帧时ARCF算法开始偏离跟踪物体的中心位置。从第32帧到第191帧时，跟踪物体被云完全遮挡且超出视野，目标物体完全丢失。在第191帧中，当目标物体再次出现时，本文算法可以继续对目标进行跟踪，而其他算法完全丢失跟踪物体，跟踪器发生漂移。在第276帧中，LADCF算法再次跟踪物体，但跟踪的物体框相比本文算法不够精确且偏离中心位置。可以直观地看出，本文算法相比其他算法具有更强的跟踪能力。第2行Skating序列中包含尺度变换、遮挡、形变、快速运动和平面外旋转。当两位滑冰运动员快速运动时，在第24帧中其他算法已经偏移跟踪中心。在第45帧中，运动员的形变导致CRSRCF和ECO算法偏移严重，引入过多的背景信息。第92帧中，本文算法仍能对目标物体进行精准跟踪。第3行Skiing序列中包含光照变化、尺度变换、形变和平面内外旋转的挑战。在第23帧中，跟踪物体受光照变化、旋转和形变的影响后，除本文算法和ECO算法，其余的4个算法都发生了漂移，失去了跟踪物体。在第36帧和第49帧中，物体尺度发生变化。可以看出，本文算法可以更好地进行尺度变化，而ECO跟踪框过大引入了过多的背景信息。第4行Biker序列中当运动发生尺度变化和变形时，在第80帧CRSRCF算法完全丢失物体，LADCF、DSARCF、ARCF也发生了不同程度的跟踪漂移。在第99帧时，对比的4个算法都丢失了物体，虽然ECO算法依旧跟踪目标物体，但是跟踪已经偏离物体中心。在第142帧中可以看出本文算法跟踪的良好表现。本文引入高斯形状掩膜改进的算法相比其他算法有更精确的预测。应对变形等挑战高斯形状掩膜有更强的能力，对于物体的突变和环境变化，在算法中引入高斯形状掩膜可以增强跟踪能力，降低边缘响应，对物体更准确地跟踪。

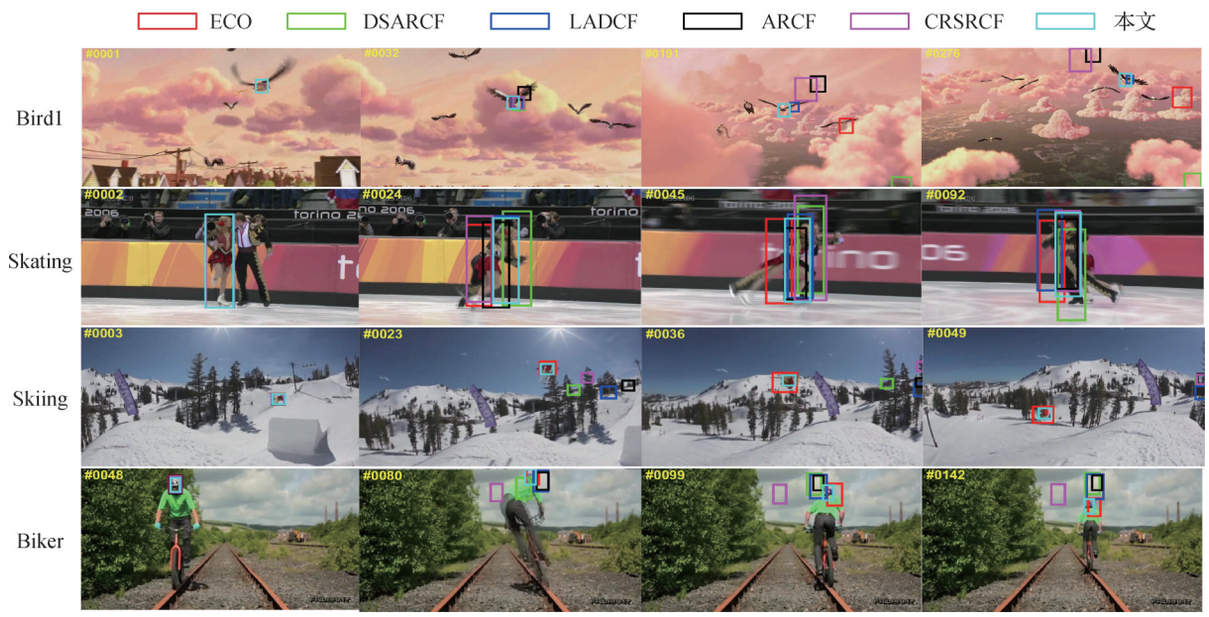


图8　在OTB2013数据集上本文算法与其他算法定性对比

Fig. 8　Qualitative comparison results of our algorithm and other methods on OTB2013 dataset

4　结　论

本文针对边界效应问题和样本污染问题，在自适应空间正则化相关滤波框架中嵌入高斯形状掩膜，对靠近目标中心和基础图像区域边缘的样本的权重重新分配，增加了中心样本的重要性，同时降低了远离中心的边缘样本的重要性，在抑制边界效应的同时降低样本污染的影响。建立了基于高斯形状掩膜的相关滤波跟踪算法模型，并通过ADMM方法优化模型，求得滤波器及空间权重的闭合解。

在OTB2013、TC128、UAV123和Got-10k等4个基准数据集上进行了大量的对比实验。在OTB2013数据集上，本文算法精度比基准算法提高0.5%；在TC128数据集上成功率提高0.4%；在UAV123数据集上精度提高0.3%；在Got-10k数据集上成功率提高了0.2%。实验结果表明，与基准算法相比，嵌入高斯掩膜提高了算法的跟踪性能。同时，与其他先

进的相关滤波跟踪算法对比也获得了有竞争力的结果。

由于本文算法嵌入了高斯形状掩膜，并且训练滤波器时使用了VGG深度特征，增加了算法的计算耗费，使得本文算法的帧率降低。另一方面，针对不同挑战属性的评估结果表明，本文算法在处理运动模糊挑战方面落后于当前先进算法。在下一步的工作中，将在本文跟踪算法框架中采用时间正则化方法，进一步提升算法的跟踪性能，同时考虑采用轻量级深度神经网络提取目标特征，降低算法的时间耗费。

参考文献（References）

Abbass M Y, Kwon K C, Kim N, Abdelwahab S A, El-Samie F E A and Khalaf A A M. 2021. A survey on online learning for visual tracking. The Visual Computer, 37(5): 993-1014 [DOI: 10.1007/s00371-020-01848-y]

Bertinetto L, Valmadre J, Golodetz S, Miksik O and Torr P H S. 2016. Staple: complementary learners for real-time tracking//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, USA: IEEE: 1401-1409 [DOI: 10.1109/CVPR.2016.156]

Bolme D S, Beveridge J R, Draper B A and Lui Y M. 2010. Visual object tracking using adaptive correlation filters//Proceedings of 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Francisco, USA: IEEE: 2544-2550 [DOI: 10.1109/CVPR.2010.5539960]

Dai K N, Wang D, Lu H C, Sun C and Li J H. 2019. Visual tracking via adaptive spatially-regularized correlation filters//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, USA: IEEE: 4665-4674 [DOI: 10.1109/CVPR.2019.00480]

Danelljan M, Bhat G, Khan F S and Felsberg M. 2017a. ECO: efficient convolution operators for tracking//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA: IEEE: 6931-6939 [DOI: 10.1109/CVPR.2017.733]

Danelljan M, Häger G, Khan F S and Felsberg M. 2015. Learning spatially regularized correlation filters for visual tracking//Proceedings of 2015 IEEE International Conference on Computer Vision (ICCV). Santiago, Chile: IEEE: 4310-4318 [DOI: 10.1109/ICCV.2015.490]

Danelljan M, Häger G, Khan F S and Felsberg M. 2017b. Discriminative scale space tracking. IEEE Transactions on Pattern Analysis and Machine Intelligence, 39(8): 1561-1575 [DOI: 10.1109/TPAMI.2016.2609928]

Elayaperumal D and Joo Y H. 2021. Aberrance suppressed spatio-temporal correlation filters for visual object tracking. Pattern Recognition, 115: #107922 [DOI: 10.1016/j.patcog.2021.107922]

Feng W, Han R Z, Guo Q, Zhu J K and Wang S. 2019. Dynamic saliency-aware regularization for correlation filter-based object tracking. IEEE Transactions on Image Processing, 28(7): 3232-3245 [DOI: 10.1109/TIP.2019.2895411]

Fu C H, Xu J T, Lin F L, Guo F Y, Liu T C and Zhang Z J. 2020. Object saliency-aware dual regularized correlation filter for real-time aerial tracking. IEEE Transactions on Geoscience and Remote Sensing, 58(12): 8940-8951 [DOI: 10.1109/TGRS.2020.2992301]

Galoogahi H K, Fagg A and Lucey S. 2017. Learning background-aware correlation filters for visual tracking//Proceedings of 2017 IEEE International Conference on Computer Vision (ICCV). Venice, Italy: IEEE: 1144-1152 [DOI: 10.1109/ICCV.2017.129]

Han R Z, Guo Q and Feng W. 2018. Content-related spatial regularization for visual object tracking//Proceedings of 2018 IEEE International Conference on Multimedia and Expo (ICME). San Diego, USA: IEEE: 1-6 [DOI: 10.1109/ICME.2018.8486487]

Huang L H, Zhao X and Huang K Q. 2021. GOT-10k: a large high-diversity benchmark for generic object tracking in the wild. IEEE Transactions on Pattern Analysis and Machine Intelligence, 43(5): 1562-1577 [DOI: 10.1109/TPAMI.2019.2957464]

Huang Z Y, Fu C H, Li Y M, Lin F L and Lu P. 2019. Learning aberrance repressed correlation filters for real-time UAV tracking//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, Korea (South): IEEE: 2891-2900 [DOI: 10.1109/ICCV.2019.00298]

Jiang W T, Tu C and Liu W J. 2021. Background and direction-aware correlation filter tracking. Journal of Image and Graphics, 26(3): 527-541 (姜文涛, 涂潮, 刘万军. 2021. 背景与方向感知的相关滤波跟踪. 中国图象图形学报, 26(3): 527-541) [DOI: 10.11834/jig.200139]

Li B W, Fu C H, Ding F Q, Ye J J and Lin F L. 2023. All-day object tracking for unmanned aerial vehicle. IEEE Transactions on Mobile Computing, 22(8): 4515-4529 [DOI: 10.1109/TMC.2022.3162892]

Li F, Tian C, Zuo W M, Zhang L and Yang M H. 2018. Learning spatial-temporal regularized correlation filters for visual tracking//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE: 4904-4913 [DOI: 10.1109/CVPR.2018.00515]

Li F, Wu X H, Zuo W M, Zhang D and Zhang L. 2020a. Remove cosine window from correlation filter-based visual trackers: when and how. IEEE Transactions on Image Processing, 29: 7045-7060 [DOI: 10.1109/TIP.2020.2997521]

Li Y M, Fu C H, Ding F Q, Huang Z Y and Lu G. 2020b. AutoTrack: towards high-performance visual tracking for UAV with automatic

spatio-temporal regularization//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, USA: IEEE: 11920-11929 [DOI: 10.1109/CVPR42600.2020.01194]

Liang P P, Blasch E and Ling H B. 2015. Encoding color information for visual tracking: algorithms and benchmark. IEEE Transactions on Image Processing, 24(12): 5630-5644 [DOI: 10.1109/TIP.2015.2482905]

Liu S, Liu D Y, Srivastava G, Połap D and Woźniak M. 2021. Overview and methods of correlation filter algorithms in object tracking. Complex and Intelligent Systems, 7(4): 1895-1917 [DOI: 10.1007/s40747-020-00161-4]

Mueller M, Smith N and Ghanem B. 2016. A benchmark and simulator for UAV tracking//Proceedings of the 14th European Conference on Computer Vision. Amsterdam, the Netherland: Springer International Publishing: 445-461 [DOI: 10.1007/978-3-319-46448-0_27]

Pu L, Feng X X, Hou Z Q, Zha Y F and Yu W S. 2020. Correlation filter tracking based on deep spatial regularization. Acta Electronica Sinica, 48(10): 2025-2032 (蒲磊, 冯新喜, 侯志强, 查宇飞, 余旺盛. 2020. 基于深度空间正则化的相关滤波跟踪算法. 电子学报, 48(10): 2025-2032) [DOI: 10.3969/j.issn.0372-2112.2020.10.021]

Qiang Z and Shi F H. 2020. Dynamic update correlation filter tracking based on appearance representation analysis. Journal of Image and Graphics, 25(6): 1209-1220 (强壮, 石繁槐. 2020. 外观表征分析下动态更新相关滤波跟踪. 中国图象图形学报, 25(6): 1209-1220) [DOI: 10.11834/jig.190454]

Qiu Z L, Zha Y F, Wu M and Wang Q. 2020. Learning attentional regularized correlation filter for visual tracking. Acta Electronica Sinica, 48(9): 1762-1768 (仇祝令, 查宇飞, 吴敏, 王青. 2020. 基于注意力学习的正则化相关滤波跟踪算法. 电子学报, 48(9): 1762-1768) [DOI: 10.3969/j.issn.0372-2112.2020.09.014]

Wu Y, Lim J and Yang M H. 2013. Online object tracking: a benchmark//Proceedings of 2013 IEEE Conference on Computer Vision and Pattern Recognition. Portland, USA: IEEE: 2411-2418 [DOI: 10.1109/CVPR.2013.312]

Xu T Y, Feng Z H, Wu X J and Kittler J. 2019. Learning adaptive discriminative correlation filters via temporal consistency preserving spatial feature selection for robust visual object tracking. IEEE Transactions on Image Processing, 28(11): 5596-5609 [DOI: 10.1109/TIP.2019.2919201]

Xu T Y, Feng Z H, Wu X J and Kittler J. 2021. Adaptive channel selection for robust visual object tracking with discriminative correlation filters. International Journal of Computer Vision, 129(5): 1359-1375 [DOI: 10.1007/s11263-021-01435-1]

Ye J J, Fu C H, Lin F L, Ding F Q, An S and Lu G. 2022. Multi-regularized correlation filter for UAV tracking and self-localization. IEEE Transactions on Industrial Electronics, 69(6): 6004-6014 [DOI: 10.1109/TIE.2021.3088366]

Yuan D, Shu X and He Z Y. 2020. TRBACF: learning temporal regularized correlation filters for high performance online visual object tracking. Journal of Visual Communication and Image Representation, 72: #102882 [DOI: 10.1016/j.jvcir.2020.102882]

Zeng M Y, Shang Z H, Liu H and Li J P. 2020. Target tracking algorithm based on adaptive updating of multilayer convolution features. Laser and Optoelectronics Progress, 57(2): #021008 (曾梦媛, 尚振宏, 刘辉, 李健鹏. 2020. 融合多层卷积特征自适应更新的目标跟踪算法. 激光与光电子学进展, 57(2): #021008) [DOI: 10.3788/LOP.021008]

Zhang Y H, Yang Y J, Zhou W N, Shi L F and Li D M. 2018. Motion-aware correlation filters for online visual tracking. Sensors, 18(11): #3937 [DOI: 10.3390/s18113937]

Zhao J W, Wei F Y, Chen N N and Zhou Z H. 2022. Spatial and long-short temporal attention correlation filters for visual tracking. IET Image Processing, 16(11): 3011-3024 [DOI: 10.1049/ipr2.12535]

Zheng G Z, Fu C H, Ye J J, Lin F L and Ding F Q. 2021. Mutation sensitive correlation filter for real-time UAV tracking with adaptive hybrid label//Proceedings of 2021 IEEE International Conference on Robotics and Automation (ICRA). Xi'an, China: IEEE: 503-509 [DOI: 10.1109/ICRA48506.2021.9561931]

作者简介

王法胜,男,教授,硕士生导师,主要研究方向为计算机视觉与模式识别。E-mail: wangfasheng@dlnu.edu.cn

孙福明,通信作者,男,教授,博士生导师,主要研究方向为机器学习和多媒体计算。E-mail: sunfuming@dlnu.edu.cn

尹双双,女,硕士研究生,主要研究方向为计算机视觉。E-mail: yss460229828@163.com

贺冰,女,硕士研究生,主要研究方向为计算机视觉。E-mail: hebing_1999@163.com