

# 自适应时空正则化的相关滤波目标跟踪

齐向明, 陈伟\*

辽宁工程技术大学软件学院, 辽宁 葫芦岛 125105

**摘要** 针对相关滤波目标跟踪算法空间正则权重没有与目标建立联系和时间正则项不能自适应更新的问题, 提出自适应时空正则化的相关滤波目标跟踪算法。首先, 利用初始帧的显著感知参考权重, 使自适应空间正则项能够在后续跟踪过程中获取与目标存在联系的空间正则权重。然后, 利用相邻两帧响应得分的变化情况计算时间正则化参数的参考值, 使自适应时间正则项可以通过变化的正则化参数不断更新。最后, 采用交替方向乘子法 (ADMM) 优化算法, 以较少的迭代次数分别求解出滤波器函数、空间正则权重和时间正则化参数。在 OTB-2015 数据集上进行实验, 结果表明本文算法的跟踪性能优于其他对比算法, 其中距离精度和成功率分别达到 86.4% 和 65.6%, 且本文算法在具有形变、旋转、遮挡和出视野等属性的复杂跟踪场景下更具鲁棒性。

**关键词** 机器视觉; 目标跟踪; 相关滤波; 自适应时空正则化; 显著感知; 响应得分

中图分类号 TP391.4

文献标志码 A

doi: 10.3788/LOP202158.0215007

## Correlation Filter Object Tracking Based on Adaptive Spatiotemporal Regularization

Qi Xiangming, Chen Wei\*

College of Software, Liaoning Technical University, Huludao, Liaoning 125105, China

**Abstract** For current correlation filter target tracking algorithm, the spatial regularization weight is not connected with an object, and the temporal regularization term fails to update adaptively. To resolve this problem, a correlation filter based on adaptive spatiotemporal regularization was proposed. The adaptive spatial regularization term first obtains the spatial regularization weight connected with the object by initial-frame saliency aware reference weight. Second, the reference value of the temporal regularization parameter is calculated using the altered response score between two adjacent frames. Thus, the adaptive temporal regularization term can be continuously updated by the changing regularization parameter. Finally, the algorithm is optimized by the alternating direction method of multipliers, which reduces the number of iterations and solves the related parameters (filtering function, spatial regularization weight, and temporal regularization parameter). In an experimental evaluation on OTB-2015 dataset, our algorithm outperformed comparable algorithms, achieving a distance precision of 86.4% and a success rate of 65.6%. The proposed algorithm also showed higher robustness in complex scenes with deformation, rotation, occlusions, and out of view than the competing algorithms.

**Key words** machine vision; object tracking; correlation filtering; adaptive spatiotemporal regularization; saliency awareness; response score

**OCIS codes** 150.1135; 100.4999; 330.7326

收稿日期: 2020-06-16; 修回日期: 2020-07-07; 录用日期: 2020-07-24

\* E-mail: 1163186035@qq.com

## 1 引言

目标跟踪是计算机视觉领域的一个重要研究分支,其融合了图像处理和最优化等理论<sup>[1]</sup>,在人机交互、行为识别和无人驾驶等<sup>[2]</sup>方面有着广泛的应用。在具有形变、旋转、遮挡和出视野等属性的复杂跟踪场景下,目标跟踪仍然是一项具有挑战性的工作。

相关滤波跟踪算法凭借其优越的跟踪精度和速度从主流跟踪方法中脱颖而出,得到了研究学者们的青睐。Bolme 等<sup>[3]</sup>在 2010 年成功地将 MOSSE (Minimum Output Sum of Squared Error) 相关滤波器应用于目标跟踪领域,随后学者们将不同的改进方法引入到相关滤波器中。Henriques 等<sup>[4]</sup>利用循环矩阵理论提出 CSK (Circulant Structure Kernels) 算法,该算法通过循环移位产生大量的训练样本,解决了 MOSSE 相关滤波器样本不足的问题。之后,他们又对 CSK 算法进行扩展,将 HOG (Histogram of Oriented Gradients) 特征和核方法加入到相关滤波器后提出 KCF (Kernel Correlation Filter) 算法<sup>[5]</sup>,此算法不仅提升了模型的判别能力,还保持了较快的跟踪速度。为解决跟踪过程中目标的尺度变化问题,国内外研究学者们在 DSST (Discriminative Scale Space Tracker) 算法<sup>[6]</sup>和 SAMF (Scale Adaptive Multiple Feature) 算法<sup>[7]</sup>的基础上进行改进,提出多种跟踪算法<sup>[8-10]</sup>。上述算法针对相关滤波器存在的问题做出了不同的改进,但边界效应和滤波器退化等问题仍然影响着相关滤波跟踪算法的性能。

Danelljian 等<sup>[11]</sup>提出空间正则化的相关滤波 (SRDCF) 跟踪算法,其通过一个固定的空间正则权重惩罚滤波器函数,使滤波器函数集中在目标的中心区域,减弱了边界效应。但空间正则权重没有与目标建立联系,在具有形变和旋转等属性的跟踪场景下,算法可能无法获取可靠的滤波器惩罚系数。Dai 等<sup>[12]</sup>提出自适应空间正则化的相关滤波 (ASRCF) 跟踪算法,算法中的自适应空间正则项在减弱边界效应的同时还可得到与目标存在联系的空间正则权重,但时间正则项没有被加入到 ASRCF 算法的目标函数中。Li 等<sup>[13]</sup>在 SRDCF 的基础上提出时空正则化的相关滤波 (STRCF) 跟踪算法,该算法利用时间正则项建立相邻两帧滤波器的联系,有效地解决了滤波器退化的问题,但时间正则项不能自适应更新,当目标受到具有遮挡和出视野等属性的跟踪场景的干扰时,滤波器一旦发生退化,则可

能会导致跟踪漂移。

本文以上述研究为基础,提出自适应时空正则化的相关滤波目标跟踪算法。在目标函数中加入显著感知自适应空间正则项,通过初始帧的显著感知参考权重,使自适应空间正则项可以在跟踪过程中获取与目标存在联系的空间正则权重,增强了滤波器惩罚系数的可靠性。之后加入响应得分自适应时间正则项,利用相邻两帧响应得分的变化情况计算时间正则化参数的参考值,使自适应时间正则项在跟踪过程中通过变化的正则化参数不断更新,避免了滤波器退化后跟踪漂移的发生。最后采用交替方向乘法 (ADMM) 进行算法优化,以较少的迭代次数分别求解出滤波器函数、空间正则权重和时间正则化参数。

## 2 相关滤波算法

### 2.1 相关滤波器

相关滤波跟踪算法可以分为训练和检测两个阶段。在训练阶段,通过当前帧的目标特征可以学习得到一个相关滤波器函数  $f \in \mathbf{R}^{M \times N \times D}$ ,其中  $M, N$  表示滤波器的大小, $D$  是从目标区域提取特征的总通道数。在检测阶段,首先利用上一帧的滤波器和新一帧的目标特征产生一个响应图  $R_t \in \mathbf{R}^{M \times N}$ ,然后计算响应图中最大值的位置即可确定新一帧的目标位置。第  $t$  帧响应图可表示为

$$R_t = \sum_{d=1}^D x_t^d * f_{t-1}^d, \quad (1)$$

式中,  $x_t^d \in \mathbf{R}^{M \times N}$  是新一帧第  $d$  通道的目标特征,  $f_{t-1}^d \in \mathbf{R}^{M \times N}$  是上一帧第  $d$  通道的滤波器函数,“\*”表示相关运算,且  $d \in \{1, 2, \dots, D\}$ 。最后,相关滤波模型要以固定的学习率更新滤波器,更新方式为

$$f_t = (1 - \eta) f_{t-1} + \eta f_t, \quad (2)$$

式中,  $f_{t-1}$  和  $f_t$  分别是第  $t-1$  帧和第  $t$  帧的滤波器函数,  $\eta$  是滤波器的学习率。为获取相关滤波器函数,将相关滤波算法的目标函数写为

$$\arg \min_f \left( \frac{1}{2} \left\| \sum_{d=1}^D x^d * f^d - y \right\|^2 + \frac{\lambda}{2} \sum_{d=1}^D \|f^d\|^2 \right), \quad (3)$$

式中,  $y \in \mathbf{R}^{M \times N}$  是理想的高斯型输出,  $\lambda$  是正则化参数。

相关滤波跟踪算法虽然已取得不错的跟踪精度和速度,但因为训练样本是由目标区域通过循环移位产生的,所以会带来边界效应。另外,由于模型在

更新阶段会对每一帧的滤波器都进行更新,因此当目标所处的跟踪场景比较复杂时,继续更新会使滤波器逐渐发生退化,最终导致跟踪漂移,从而影响算法的跟踪性能。

## 2.2 空间正则化

为解决边界效应,SRDCF 算法在相关滤波算法的目标函数中加入了空间正则项,提升了算法的跟踪性能。空间正则化的相关滤波跟踪算法的目标函数为

$$\arg \min_f \left( \frac{1}{2} \left\| \sum_{d=1}^D \mathbf{x}_t^d * \mathbf{f}^d - \mathbf{y} \right\|^2 + \frac{\lambda}{2} \sum_{d=1}^D \|\mathbf{W} \odot \mathbf{f}^d\|^2 \right), \quad (4)$$

式中,  $E_{\text{SR}} = \frac{\lambda}{2} \sum_{d=1}^D \|\mathbf{W} \odot \mathbf{f}^d\|^2$  是空间正则项,  $\mathbf{W} \in \mathbf{R}^{M \times N}$  是空间正则权重,“ $\odot$ ”表示对应元素相乘。SRDCF 算法利用一个固定的空间正则权重惩罚滤波器函数,使滤波器函数集中在目标的中心区域,减小了边界效应带来的影响,则空间正则权重  $\mathbf{W}$  表示为

$$W(x, y) = a \left( \frac{x - M/2}{w/2} \right)^2 + a \left( \frac{y - N/2}{h/2} \right)^2 + b, \quad (5)$$

式中:  $x=1, 2, \dots, M$ ;  $y=1, 2, \dots, N$ ;  $w$  和  $h$  分别是目标的宽和高;  $a$  和  $b$  都是固定参数。

## 2.3 时间正则化

为解决滤波器的退化问题,STRCF 算法在 SRDCF 的基础上,利用加入到目标函数的时间正则项,进一步提升了算法的跟踪性能。时空正则化的相关滤波跟踪算法将目标函数写为

$$\arg \min_f \left( \frac{1}{2} \left\| \sum_{d=1}^D \mathbf{x}_t^d * \mathbf{f}^d - \mathbf{y} \right\|^2 + \frac{1}{2} \sum_{d=1}^D \|\mathbf{W} \odot \mathbf{f}^d\|^2 + \frac{\mu}{2} \sum_{d=1}^D \|\mathbf{f}^d - \mathbf{f}_{t-1}^d\|^2 \right), \quad (6)$$

式中,  $E_{\text{TR}} = \frac{\mu}{2} \sum_{d=1}^D \|\mathbf{f}^d - \mathbf{f}_{t-1}^d\|^2$  是时间正则项,  $\mu$  是时间正则化参数,  $\mathbf{f}_{t-1}$  是上一帧学习得到的滤波器。STRCF 通过时间正则项建立了相邻两帧滤波器之间的联系,有效地解决了滤波器退化的问题。

# 3 自适应时空正则化的相关滤波算法

## 3.1 显著感知自适应空间正则化

由于 SRDCF 算法和 STRCF 算法中的空间正则权重不能自适应更新,为此提出显著感知自适应

空间正则项。通过初始帧的显著感知参考权重,自适应空间项在跟踪过程中能够获取与目标存在联系的空间正则权重。

首先,利用显著性检测算法<sup>[14]</sup>得到第一帧图像中目标的显著性检测结果  $\mathbf{S}$ ,并对  $\mathbf{S}$  进行以下处理:

$$\mathbf{S}'(x, y) = \frac{\max \mathbf{S} - \mathbf{S}(x, y)}{\max \mathbf{S} - \min \mathbf{S}}, \quad (7)$$

式中,  $\max \mathbf{S}$  和  $\min \mathbf{S}$  分别表示显著性检测结果中的最大值和最小值。经过(7)式处理后的显著性检测结果  $\mathbf{S}'$  取值在 0 到 1 之间,即  $\mathbf{S}'(x, y) \in [0, 1]$ ,靠近目标中心区域的  $\mathbf{S}'$  取值接近于 0,周围非目标区域的  $\mathbf{S}'$  取值接近于 1。然后将处理后的显著性检测结果  $\mathbf{S}'$  裁剪成与空间正则权重  $\mathbf{W}$  大小相同的显著性图  $\mathbf{S}'' \in \mathbf{R}^{M \times N}$ 。最后,建立初始帧的显著感知参考权重  $\boldsymbol{\phi}^r$ ,表达式为

$$\boldsymbol{\phi}^r = \mathbf{W} \odot \mathbf{S}'', \quad (8)$$

式中,  $\mathbf{W}$  是通过(5)式构造获得,其中固定参数  $a=3$ ,  $b=0.1$ 。将初始帧的显著感知参考权重  $\boldsymbol{\phi}^r$  加入到自适应空间正则项后可得到显著感知自适应空间正则项,其形式为

$$E_{\text{SAASR}} = \frac{\lambda_1}{2} \sum_{d=1}^D \|\boldsymbol{\phi} \odot \mathbf{f}^d\|^2 + \frac{\lambda_2}{2} \|\boldsymbol{\phi} - \boldsymbol{\phi}^r\|^2, \quad (9)$$

式中,  $\boldsymbol{\phi}$  是与目标存在联系的空间正则权重,  $\boldsymbol{\phi}^r$  是初始帧的显著感知参考权重,  $\lambda_1$  和  $\lambda_2$  是显著感知自适应空间正则化参数。图 1 是 Soccer 视频序列通过第一帧图像获取的初始帧的显著感知参考权重。

在得到初始帧的显著感知参考权重后,可利用自适应空间正则项获取与目标存在联系的空间正则权重。为体现空间正则权重的不同,给出算法在跟踪过程中 4 种空间正则权重的可视化结果,如图 2 所示,其中:  $\mathbf{W}_1$  是 SRDCF 算法中的空间正则权重,其根据目标的中心位置对周围的背景进行惩罚,离目标越远惩罚系数越大,从而使滤波器集中在目标的中心区域;  $\mathbf{W}_2$  是 STRCF 算法中的空间正则权重,利用该算法的空间正则权重直接对目标周围的背景进行相同的惩罚,可以很好地区分目标与背景;  $\mathbf{W}_3$  是利用目标显著特征构造的空间正则权重,在跟踪过程中,由于每一帧目标的显著特征都不同,因此构造出的空间正则权重可以根据目标的大小变化对背景进行惩罚;  $\mathbf{W}_4$  是与目标存在联系的空间正则权重,在构造权重的过程中,将自适应空间正则项和显著特征相融合,通过建立显著感知的空间参考权重,使自适应空间正则项可以在跟踪过程中获取与目标



图 1 Soccer 视频序列获取初始帧的显著感知参考权重过程。(a)第一帧图像;(b)显著性检测结果;  
(c)处理后的显著性检测结果;(d)显著性图;(e)初始帧的显著感知参考权重

Fig. 1 Process of obtaining initial frame saliency awareness reference weight for Soccer video sequence. (a) First frame image; (b) saliency detection result; (c) result of saliency detection after treatment; (d) saliency image; (e) saliency awareness reference weight for initial frame

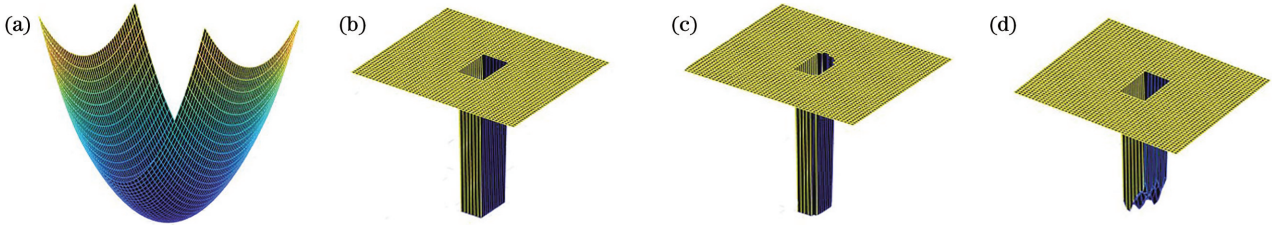


图 2 四种不同空间正则权重的可视化结果。(a) $W_1$ ;(b) $W_2$ ;(c) $W_3$ ;(d) $W_4$

Fig. 2 Visualization of four different spatial regularization weights. (a)  $W_1$ ; (b)  $W_2$ ; (c)  $W_3$ ; (d)  $W_4$

存在联系的权重,其不仅随着目标的变化不断地更新,还带有目标的显著信息。

利用第一帧图像目标的显著性检测结果和固定的空间正则权重,得到了初始帧的显著感知参考权重。之后在整个跟踪过程中,利用自适应空间正则项获取与目标存在联系的空间正则权重,在目标发生形变和旋转时,算法仍然可以获取准确的空间正则权重,从而增强滤波器惩罚系数的可靠性。

### 3.2 响应得分自适应时间正则化

在 STRCF 算法中,时间正则项的正则化参数是一个固定值,其在跟踪过程中保持不变。本文为得到自适应时间正则项,从时间正则化参数入手,根据相邻两帧响应得分的变化情况获取时间正则化参数的参考值,并在算法优化时利用变化的正则化参数自适应更新时间正则项。

首先,将相邻两帧的响应得分及其变化情况定义为

$$\begin{cases} s = \exp(-\|\mathbf{R}' - \mathbf{R}\|^2) \\ \Delta s = s_t - s_{t-1} \end{cases}, \quad (10)$$

式中, $\mathbf{R}'$ 是理想响应, $\mathbf{R}$ 是当前帧的响应, $s_t$ 和 $s_{t-1}$ 分别是第 $t$ 帧和第 $t-1$ 帧的响应得分, $\Delta s$ 是响应得分的变化情况。然后通过相邻两帧响应得分的变化情况计算时间正则化参数的参考值,即

$$\rho^r = \frac{\varphi}{1 + \kappa \Delta s}, \quad (11)$$

式中, $\varphi$ 和 $\kappa$ 是固定参数, $\rho^r$ 是时间正则化参数的参考值。当相邻两帧响应得分的变化情况 $\Delta s$ 小于阈值 $\theta$ 时,当前帧的响应图可能发生变化,此时利用(11)式获得一个时间正则化参数的参考值,并通过此参考值得到最终的时间正则化参数,使滤波器迅速学习到变化的信息;当 $\Delta s$ 大于阈值 $\theta$ 时,当前帧的响应图异常,目标可能受到遮挡和出视野等干扰,此时停止模型更新,以防止滤波器学习到错误的干扰信息。图 3 是 Box 视频序列中目标在不同情况下的响应图,其中:在第 130 帧时,目标处于较好的跟踪环境,没有受到其他物体的干扰,只是目标本身发生微小变化,此时响应图正常;在第 460 帧时,目标的跟踪环境发生变化,目标被其他物体遮挡并受到背景的干扰,目标的有效信息减少,响应图发生剧烈变化。如果发生严重的遮挡,还将导致算法无法确定目标的准确位置,此时就需要控制滤波器的学习能力,以防止其学习到错误的信息。

最后,将时间正则化参数的参考值和不断变化的正则化参数加入到时间正则项,即可得到响应得分的自适应时间正则项,表达式为

$$E_{\text{RSATR}} = \frac{\rho}{2} \sum_{d=1}^D \|\mathbf{f}^d - \mathbf{f}_{t-1}^d\|^2 + \frac{1}{2} |\rho - \rho^r|^2, \quad (12)$$

式中, $\rho$ 是可以自适应更新的时间正则化参数。通过相邻两帧响应得分的变化情况,获取时间正则化

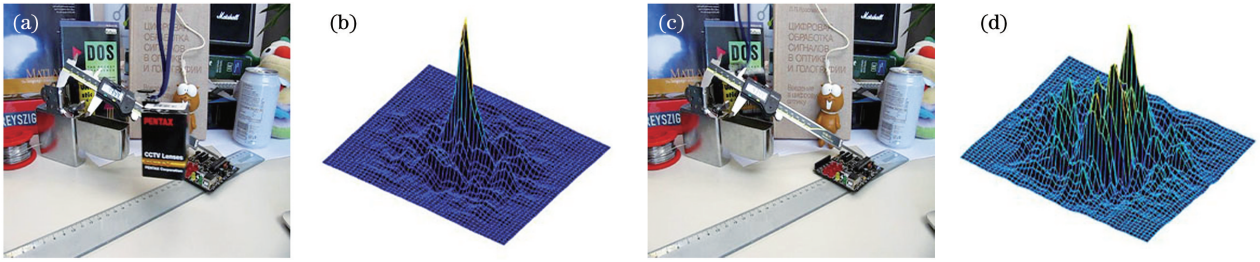


图 3 Box 视频序列中目标在不同情况下的响应图。(a)第 130 帧的目标;(b)第 130 帧的响应图;  
(c)第 460 帧的目标;(d)第 460 帧的响应图

Fig. 3 Response of object in different cases for Box video sequence. (a) Object for 130<sup>th</sup> frame; (b) response for 130<sup>th</sup> frame; (c) object for 460<sup>th</sup> frame; (d) response for 460<sup>th</sup> frame

参数的参考值,在跟踪过程中,时间正则化参数可根据参考值的变化,自适应更新时间正则项,当目标处于具有遮挡和出视野等属性的复杂跟踪场景时,算法可以避免滤波器退化引起的跟踪漂移。

综合考虑显著感知自适应空间正则项和响应得分自适应时间正则项,将自适应时空正则化的相关滤波目标跟踪算法的目标函数写为

$$\arg \min_{f, \phi, \rho} \left( \frac{1}{2} \sum_{d=1}^D \|x^d * f^d - y\|^2 + \frac{\lambda_1}{2} \sum_{d=1}^D \|\phi \odot f^d\|^2 + \frac{\lambda_2}{2} \|\phi - \phi^r\|^2 + \frac{\rho}{2} \sum_{d=1}^D \|f^d - f_{t-1}^d\|^2 + \frac{1}{2} |\rho - \rho^r|^2 \right), \quad (13)$$

式中,第 2 项和第 3 项是显著感知自适应空间正则项  $E_{SAASR}$ ,第 4 项和第 5 项是响应得分自适应时间正则项  $E_{RSATR}$ 。

### 3.3 滤波器函数和时间正则化参数的求解

利用 ADMM 算法获取(13)式的全局最优解,为使优化过程简便,本研究在求解滤波器函数、空间正则权重和时间正则化参数时,假设特征的总通道数  $D=1$ 。因此,(13)式可以写为

$$\arg \min_{f, \phi, \rho} \left( \frac{1}{2} \|x * f - y\|^2 + \frac{\lambda_1}{2} \|\phi \odot f\|^2 + \frac{\lambda_2}{2} \|\phi - \phi^r\|^2 + \frac{\rho}{2} \|f - f_{t-1}\|^2 + \frac{1}{2} |\rho - \rho^r|^2 \right). \quad (14)$$

因为目标函数中加入了响应得分自适应时间正则项,所以算法优化的过程中不仅要求解滤波器函数,还要将不断变化的时间正则化参数求解出来,从而实现自适应时间正则化。为了使用 ADMM 优化算法,首先引入一个辅助变量  $g = f$ ,然后将(14)式写成增广拉格朗日形式,则

$$L(g, \phi, \rho, v) = \frac{1}{2} \|x * f - y\|^2 + \frac{\lambda_1}{2} \|\phi \odot g\|^2 + \frac{\lambda_2}{2} \|\phi - \phi^r\|^2 + (f - g)^T v + \frac{\gamma}{2} \|f - g\|^2 + \frac{\rho}{2} \|f - f_{t-1}\|^2 + \frac{1}{2} |\rho - \rho^r|^2, \quad (15)$$

式中, $v$  和  $\gamma$  分别是拉格朗日乘子和步进参数。令  $h = \frac{1}{\gamma} v$ ,将(15)式重写为

$$L(g, \phi, \rho, h) = \frac{1}{2} \|x * f - y\|^2 + \frac{\lambda_1}{2} \|\phi \odot g\|^2 + \frac{\lambda_2}{2} \|\phi - \phi^r\|^2 + \frac{\gamma}{2} \|f - g + h\|^2 + \frac{\rho}{2} \|f - f_{t-1}\|^2 + \frac{1}{2} |\rho - \rho^r|^2, \quad (16)$$

(16)式可以通过 ADMM 算法迭代求解子问题  $f, g, \rho$ ,表达式为

$$\begin{cases} f^{i+1} = \arg \min_f (\|x * f - y\|^2 + |\rho| \|f - f_{t-1}\|^2 + \gamma \|f - g + h\|^2) \\ g^{i+1} = \arg \min_g (\|\phi \odot g\|^2 + \gamma \|f - g + h\|^2) \\ \rho^{(i+1)} = \arg \min_\rho (\rho \|f - f_{t-1}\|^2 + |\rho - \rho^r|^2) \\ h^{i+1} = h^i + f^{i+1} - g^{i+1} \end{cases}. \quad (17)$$

通过求解每个子问题的闭式解,即可得到目标函数的最优解,详细的求解过程如下。

子问题  $f$ : 利用帕塞瓦尔定理将 (17) 式中的第 1 个子问题写为

$$\arg \min_f \left( \|\hat{x} \cdot \hat{f} - \hat{y}\|^2 + \rho \|\hat{f} - \hat{f}_{t-1}\|^2 + \gamma \|\hat{f} - \hat{g} + \hat{h}\|^2 \right), \quad (18)$$

式中,“ $\cdot$ ”表示变量的离散傅里叶变换(DFT)。对 (18) 式求导,令其导数为 0,则可求出

$$\hat{f} = [\hat{x}\hat{x}^T + (\rho + \gamma)\mathbf{I}]^{-1} (\hat{x}\hat{y} + \gamma\hat{g} - \gamma\hat{h} + \rho\hat{f}_{t-1}), \quad (19)$$

利用 Sherman-Morrison 公式取代 (19) 式中的逆运算,将 (19) 式写为

$$\hat{f} = \frac{1}{\gamma + \rho} \left( \mathbf{I} - \frac{\hat{x}\hat{x}^T}{\gamma + \rho + \hat{x}^T\hat{x}} \right) \times (\hat{x}\hat{y} + \gamma\hat{g} - \gamma\hat{h} + \rho\hat{f}_{t-1}). \quad (20)$$

(20) 式中只包含乘法和加法运算,所以  $\hat{f}$  的计算相对简单。最终可通过求取  $\hat{f}$  的离散傅里叶逆变换(IDFT)得到  $f$  的值。

子问题  $g$ : 因为 (17) 式中的第 2 个子问题不包含相关运算,所以直接对其求导,令导数为 0,则

$$\mathbf{g} = (\boldsymbol{\phi}^T \boldsymbol{\phi} + \gamma \mathbf{I})^{-1} (\gamma \mathbf{f} + \gamma \mathbf{h}). \quad (21)$$

子问题  $\rho$ : 在获得 (17) 式中  $\hat{f}$  和  $\hat{f}_{t-1}$  的值后,可直接求解 (17) 式中的第 3 个子问题,则  $\rho$  的求解公式如下:

$$\rho = \rho^r - \frac{\|\hat{f} - \hat{f}_{t-1}\|^2}{2}. \quad (22)$$

更新步进参数  $\gamma$ : 步进参数  $\gamma$  的更新方式如下。

$$\gamma^{(i+1)} = \min(\gamma^{\max}, \alpha \gamma^{(i)}), \quad (23)$$

式中,  $\gamma^{\max}$  是步进参数的最大值,  $\alpha$  是固定参数。

### 3.4 空间正则权重的求解

在算法的目标函数中加入显著感知自适应空间正则项后,为得到准确且可靠的空间正则权重,使用另一个 ADMM 算法求解空间正则权重。将 (16) 式中包含空间正则权重的部分写为

$$\arg \min_{\boldsymbol{\phi}} (\lambda_1 \|\boldsymbol{\phi} \odot \mathbf{g}\|^2 + \lambda_2 \|\boldsymbol{\phi} - \boldsymbol{\phi}^r\|^2). \quad (24)$$

步骤同上,首先引入辅助变量  $\mathbf{q} = \boldsymbol{\phi}$ ,然后将 (24) 式写成增广拉格朗日形式,即

$$L(\mathbf{q}, \boldsymbol{\phi}, \mathbf{m}) = \frac{\lambda_1}{2} \|\boldsymbol{\phi} \odot \mathbf{g}\|^2 + \frac{\lambda_2}{2} \|\mathbf{q} - \boldsymbol{\phi}^r\|^2 + \frac{\zeta}{2} \|\boldsymbol{\phi} - \mathbf{q} + \mathbf{m}\|^2, \quad (25)$$

式中,  $\zeta$  和  $\mathbf{m}$  分别是步进参数和拉格朗日乘子。将

(25) 式分为 3 个子问题,通过 ADMM 算法迭代求解每个子问题即可获得最终的空间正则权重。

子问题  $\boldsymbol{\phi}$ : 通过 (21) 式得到辅助变量  $\mathbf{g}$  的值后,即可求得  $\boldsymbol{\phi}$ , 则

$$\boldsymbol{\phi} = (\mathbf{q} - \mathbf{m}) \left( \mathbf{I} + \frac{\lambda_1}{\zeta} \mathbf{g} \odot \mathbf{g} \right)^{-1}. \quad (26)$$

子问题  $\mathbf{q}$ : 辅助变量  $\mathbf{q}$  的求解公式为

$$\mathbf{q} = \frac{\lambda_2 \boldsymbol{\phi}^r + \zeta (\boldsymbol{\phi} + \mathbf{m})}{\lambda_2 + \zeta}. \quad (27)$$

子问题  $\mathbf{m}$ : 在获得  $\boldsymbol{\phi}$  和  $\mathbf{q}$  的值后,可求解拉格朗日乘子,公式为

$$\mathbf{m}^{i+1} = \mathbf{m}^i + \boldsymbol{\phi}^{i+1} - \mathbf{q}^{i+1}, \quad (28)$$

式中,  $\boldsymbol{\phi}^{i+1}$  和  $\mathbf{q}^{i+1}$  分别是子问题  $\boldsymbol{\phi}$  和子问题  $\mathbf{q}$  迭代  $i+1$  次得到的值。步进参数  $\zeta$  以  $\zeta^{(i+1)} = \min(\zeta^{\max}, \beta \zeta^{(i)})$  的方式更新,其中  $\zeta^{\max}$  是步进参数  $\zeta$  的最大值,  $\beta$  是固定参数。

在算法优化的整个过程中,将滤波器函数、空间正则权重和时间正则化参数的求解分成两个部分,第 1 部分是求解滤波器函数和时间正则化参数,第 2 部分是单独求解空间正则权重。首先,通过第 1 部分的 (21) 式获得辅助变量  $\mathbf{g}$  的值,然后将其代入第 2 部分的 (26) 式求得空间正则权重  $\boldsymbol{\phi}$ ,最后直接利用空间正则权重  $\boldsymbol{\phi}$  求出最后的滤波器函数  $\mathbf{f}$ ,即计算出响应的滤波器函数。在完成第 1 部分和第 2 部分的求解后,算法在模型更新阶段更新显著感知参考权重和理想响应,更新方式分别是  $\boldsymbol{\phi}^r = \boldsymbol{\phi}$  和  $\mathbf{R}' = (1 - \delta)\mathbf{R}' + \delta\mathbf{R}$ ,其中  $\delta$  为理想响应的学习率。

## 4 实验与结果分析

### 4.1 实验环境与参数设置

本实验都在 Intel i5 6900 CPU 8 GB 内存的计算机上进行,其中操作系统为 Window 10,软件为 Matlab2017a。自适应时空正则化的相关滤波目标跟踪算法使用 HOG、CN 和灰度三种特征,相关参数设置与 STRCF 相同。尺度估计阶段,利用 SAMF 算法中的尺度池搜索策略获取目标最优尺度,尺度个数为 5,尺度步长为 1.01。在显著感知自适应空间正则项和响应得分自适应时间正则项中,根据文献[13]和文献[15]中已有的参数设置,将  $\lambda_1$  和  $\lambda_2$  分别设置为 1.00 和 0.98,固定参数  $\varphi$  和  $\kappa$  分别设置为 16 和 1,阈值  $\theta$  为 0.3。算法优化过程中,利用 ADMM 算法分别求解滤波器函数、时间正则化参数和空间正则权重。在求解滤波器函数和时间

正则化参数时, ADMM 方法的相关参数设置与 STRCF 相同; 求解空间正则权重时, 参考 ASRCF 算法, 选取 4 组不同的参数测试本文算法的跟踪性能, 并通过对比得出最终的参数设置。本文算法相

关参数的设置如表 1 所示, 其中  $N_1$  和  $N'_1$  表示 ADMM 算法的迭代次数,  $\gamma^{(0)}$  和  $\zeta^{(0)}$  是步进参数的初始值。在模型更新阶段, 理想响应的学习率  $\delta=0.018$ 。

表 1 算法优化过程中 ADMM 的相关参数设置

Table 1 Relevant parameters setting of ADMM in the process of algorithm optimization

ADMM	Relevant parameter
Solving filtering function and temporal regularization parameter	$\gamma^{(0)}=1, \gamma^{\max}=10, \alpha=0.1, N_1=2$
Solving spatial regularization weight	$\zeta^{(0)}=1, \zeta^{\max}=100, \beta=10, N'_1=2$

### 4.2 数据集和评估标准

实验采用 OTB-2015 数据集<sup>[16]</sup>, 此数据集有 100 个不同的视频序列, 且每个视频序列都具有形变、旋转、遮挡和出视野等属性的多种跟踪场景, 可满足测试算法跟踪性能的要求。实验过程中, 采用一次通过评估标准(OPE), 并使用两个指标评估算法: 距离精度(DP)和成功率(SR)。其中: 距离精度是算法得到的目标位置和实际位置之间的平均欧氏距离小于阈值( $t_d=20$  pixel)的帧数与全部帧数的百分比; 成功率是算法得到的目标跟踪框和实际跟踪框之间的重叠率大于阈值( $t_o=0.5$ )的帧数与全部帧数的百分比。

### 4.3 对比实验与结果分析

在算法优化过程中, 利用 ADMM 方法求解空间正则权重。为以较少的迭代次数求解出最优的空间正则权重, 根据 ASRCF 算法已有的参数设置, 在保持文中其他参数一致的情况下, 选取 4 组不同的参数测试算法在 OTB-2015 数据集上的跟踪性能, 通过对比距离精度和成功率, 得出空间正则权重的

最优参数。对比实验结果如表 2 所示, 且表中加粗字体为最优值。

表 2 不同参数设置下算法的跟踪性能对比

Table 2 Comparison of tracking performance of algorithm in different parameter settings unit: %

Parameter setting	DP	SR
$\zeta^{(0)}=1, \zeta^{\max}=1000, \beta=10, N'_1=2$	86.2	65.1
$\zeta^{(0)}=1, \zeta^{\max}=10000, \beta=10, N'_1=2$	84.1	63.8
$\zeta^{(0)}=1, \zeta^{\max}=100, \beta=10, N'_1=2$	<b>86.4</b>	<b>65.6</b>
$\zeta^{(0)}=1, \zeta^{\max}=100, \beta=10, N'_1=3$	84.9	64.3

通过对比实验得出本文算法最终的参数设置后, 为检验自适应时空正则化的相关滤波目标跟踪算法的跟踪性能, 采用 OTB-2015 数据集, 选取 7 种跟踪算法与本文算法作比较, 其中包括 6 种基于相关滤波的跟踪算法 (BACF<sup>[17]</sup>、DSST<sup>[6]</sup>、KCF<sup>[5]</sup>、SAMF<sup>[7]</sup>、SRDCF<sup>[11]</sup>、STRCF<sup>[13]</sup>) 和 1 种深度学习跟踪算法 (CNN-SVM<sup>[18]</sup>), 实验结果如图 4 所示。

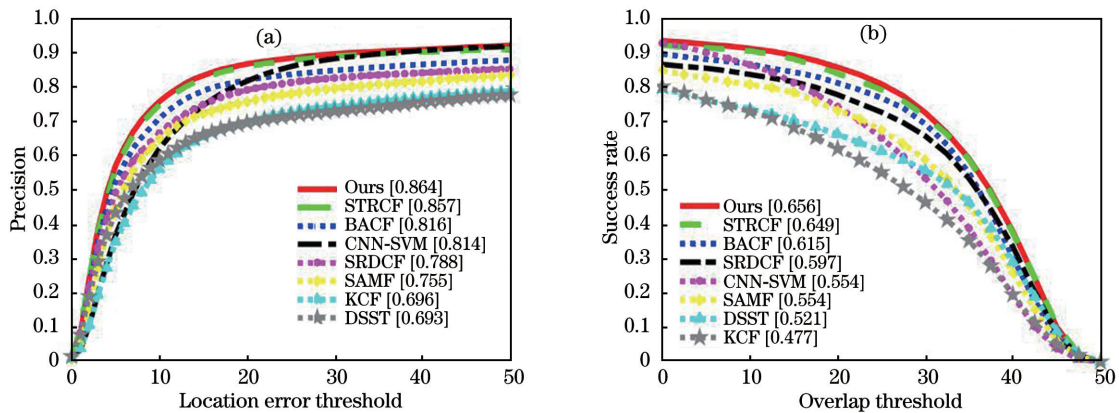


图 4 OTB-2015 数据集上的距离精度和成功率。(a) 距离精度曲线; (b) 成功率曲线

Fig. 4 Distance precision and success rate on OTB-2015 dataset. (a) Distance precision curve; (b) success rate curve

分析 8 种跟踪算法在 OTB-2015 数据集上的距离精度曲线和成功率曲线, 可知本文算法在

OTB-2015 数据集上的跟踪性能最优, 其距离精度和成功率分别达到 86.4% 和 65.6%。与 SRDCF

算法相比,距离精度和成功率分别提高 7.6 个百分点和 5.9 个百分点;与 STRCF 算法相比,距离精度和成功率都提高了 0.7 个百分点;与深度学习跟踪算法 CNN-SVM 相比,本文算法仍然可以取得较优的跟踪效果,距离精度和成功率分别提高 5.0 个百分点和 10.2 个百分点。在跟踪速度

方面,由于算法优化过程中对滤波器部分进行了两次求解,所以速度较 STRCF 有所下降,但与 SRDCF 和 CNN-SVM 相比,速度较快,可满足实时性要求。表 3 给出本文算法与 DSST、BACF、SRDCF、STRCF 和 CNN-SVM 在 OTB-2015 数据集上的平均跟踪速度对比。

表 3 OTB-2015 数据集上的平均跟踪速度对比

Table 3 Comparison of average tracking speeds on OTB-2015 dataset

Algorithm	DSST	BACF	SRDCF	STRCF	CNN-SVM	Ours
Average tracking speed / (frame · s <sup>-1</sup> )	<b>31.8</b>	22.1	3.3	13.7	N	12.1

虽然本文算法在 OTB-2015 数据集上的成功率与 STRCF 相比只提高了 0.7 个百分点,但在目标函数中加入显著感知自适应空间正则项和自适应时间正则项后,算法在具有形变、旋转、遮挡和出视野等属性的复杂跟踪场景下的跟踪性能有较为明显的提升。图 5 是自适应时空正则化的相关滤波目标跟踪算法与对比算法在 OTB-2015 数据集 4 种不同属性视频序列上的成功率曲线。由图 5 可以看出,本文算法在具有形变和旋转(平面

内旋转)属性的视频序列上,成功率分别 61.9% 和 63.2%,在具有遮挡和出视野属性的视频序列上,成功率分别为 62.1% 和 59.8%。与 SRDCF 和 STRCF 在这 4 种不同属性视频序列上的成功率相比,本文算法的成功率为最优值,这说明提出的显著感知自适应空间正则项可以使算法在跟踪过程中获取与目标存在联系的空间正则权重,从而对滤波器惩罚系数进行可靠的惩罚,最终使算法在具有形变和旋转等属性的跟踪场景下的鲁棒性得

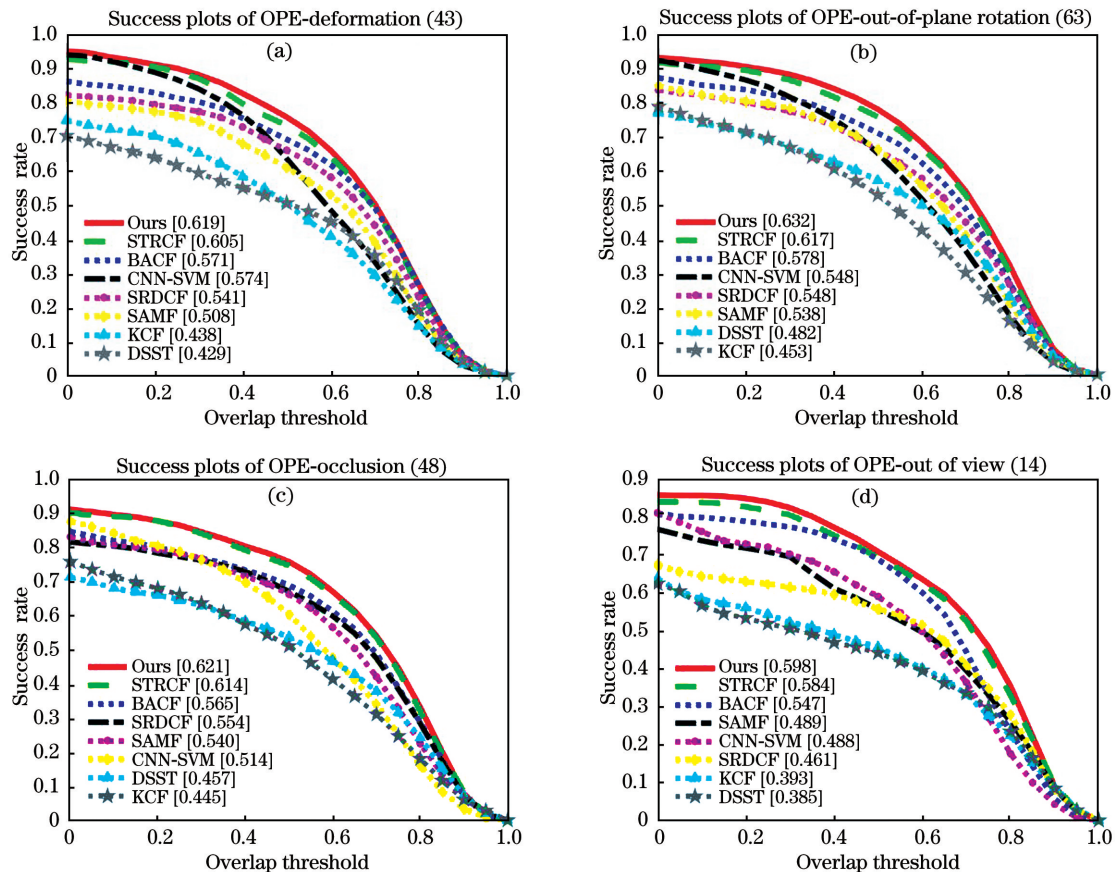


图 5 OTB-2015 数据集 4 种不同属性视频序列上的成功率。(a)形变;(b)平面外旋转;(c)遮挡;(d)出视野

Fig. 5 Success rate of 4 different attribute video sequences on OTB-2015 dataset. (a) Deformation;

(b) out-of-plane rotation; (c) occlusion; (d) out of view



以提升。自适应时间正则项加入到目标函数后,算法利用相邻两帧响应得分的变化情况得到了时间正则化参数的参考值,在参考值不断变化的情况下,时间正则项可以根据跟踪结果的好坏来获取正则化参数,从而控制滤波器的更新,克服了具有遮挡、出视野等属性的跟踪场景对算法的干扰,

提升了算法的跟踪性能。

最后,本文从 OTB-2015 数据集中挑选 5 个具有代表性的视频序列对算法进行定性分析。图 6 是 BACF、CNN-SVM、SAMF、SRDCF 和本文算法在视频序列 Box、DragonBaby、Shaking 和 Soccer 上的跟踪结果。

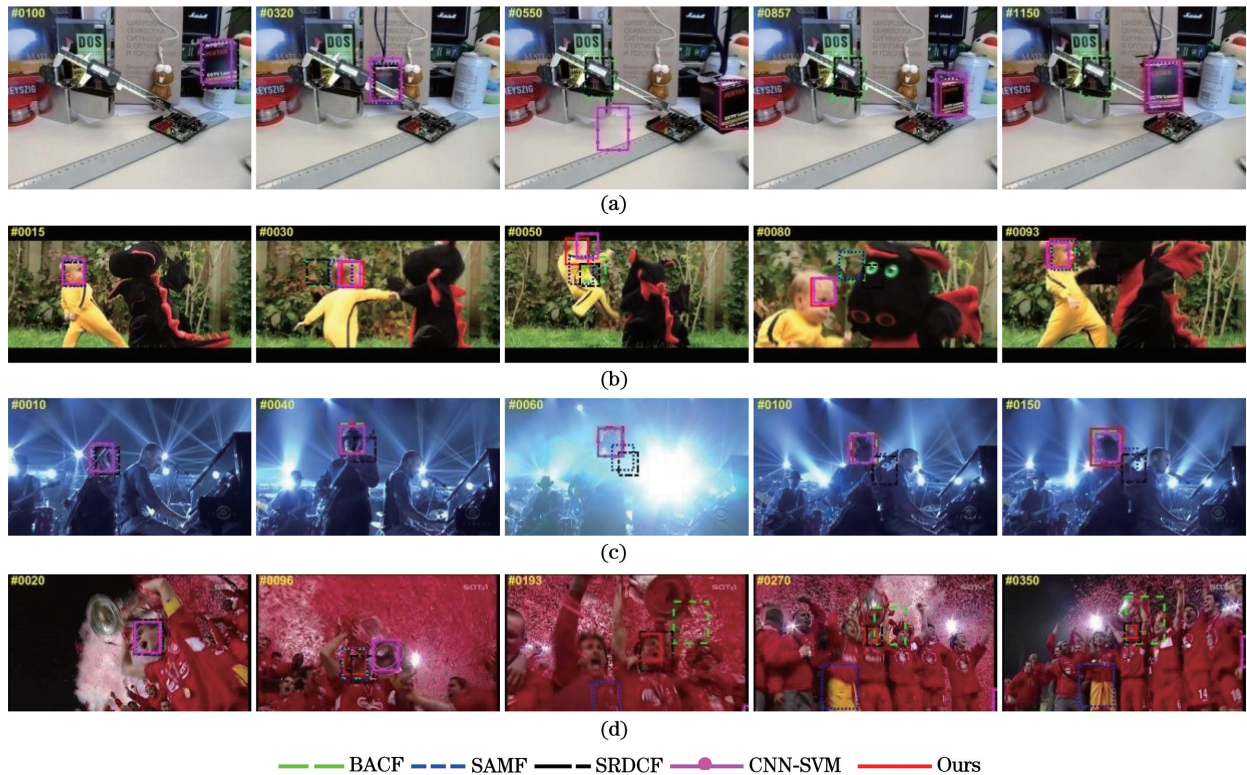


图 6 本文算法与对比算法在 4 个视频序列上的跟踪结果。(a)Box;(b)dragonBaby;(c)shaking;(d)soccer

Fig. 6 Tracking results of our algorithm and comparison algorithms on 4 video sequences. (a) Box; (b) dragonbaby; (c) shaking; (d) soccer

Box 视频序列具有遮挡、旋转和背景杂乱等属性。在第 1 帧到第 100 帧的这段跟踪过程中,目标发生旋转和尺度变化后,5 种算法仍然可以较好地跟踪目标,但 CNN-SVM 不能适应目标的尺度变化;在第 320 帧时,目标受到部分遮挡和背景杂乱的影响,CNN-SVM 的跟踪框发生偏移,其余算法都能较好地处理周围信息的干扰;第 550 帧对应目标被完全遮挡之后的情况,BACF 和 SRDCF 的跟踪框停留在遮挡物上,CNN-SVM 失去跟踪目标,只有 SAMF 和本文算法可以确定目标位置。在第 857 帧时,使用深度特征的 CNN-SVM 又重新捕获到目标并进行稳定跟踪;在第 1150 帧时,SAMF、CNN-SVM 和本文算法可以准确跟踪目标,而 BACF 和 SRDCF 一直跟踪之前的遮挡物,无法重新确定目标的位置。

DragonBaby 视频序列具有遮挡、模糊、快速运

动和旋转等属性。在第 15 帧时,本文算法和对比算法都可以较好地跟踪目标;目标在第 30 帧时发生了旋转和快速运动,此时 BACF 和 SRDCF 失去跟踪目标,其余算法仍能确定目标位置;在第 50 帧时,只有本文算法可以较好地确定目标的位置和尺度,其余算法均受到目标快速运动、模糊和旋转的影响。第 80 帧对应目标发生遮挡和旋转之后各种算法的跟踪结果,SRDCF 跟踪到另一个目标,BACF 和 SAMF 的跟踪框停留在背景上,CNN-SVM 和本文算法的跟踪框发生了变化,但仍然能确定目标位置;在第 93 帧时,SRDCF 仍然不能跟踪到正确的目标,本文算法和其他对比算法可以稳定地跟踪目标。

Shaking 视频序列具有光照变化、形变和旋转等属性。在第 10 帧时,CNN-SVM 可以较好地跟踪目标,其他算法的跟踪框都发生了偏移;在第 40 帧

时,目标发生旋转,SAMF 和 SRDCF 的跟踪框移动到目标周围的背景上,BACF、CNN-SVM 和本文算法依然可以确定目标的位置;第 60 帧对应目标受到光照变化影响后的情况,此时 SRDCF 和 SAMF 仍然不能跟踪到目标,本文算法、BACF 和 CNN-SVM 克服了强烈的光照变化影响,可以准确地跟踪目标;从第 100 帧到第 150 帧,目标在跟踪过程中发生了旋转、尺度变化等,除了 SAMF 和 SRDCF 不能确定目标外,其余算法都可以一直稳定地跟踪目标。

Soccer 视频序列具有遮挡、模糊、旋转背景杂乱等属性。第 20 帧时目标还没有受到包含遮挡与背景杂乱等属性的跟踪场景的干扰,5 种算法的跟踪框都可准确确定目标的位置;但从第 96 帧开始,目标受到相似遮挡物以及旋转的影响,CNN-SVM 首先丢失跟踪目标,其跟踪框出现在其他干扰物上,SRDCF 的跟踪框也发生了部分偏移;在第 193 帧时,只有本文算法和 SRDCF 还能跟踪到目标,BACF 的跟踪框出现与目标相似的遮挡物上,SAMF 和 CNN-SVM 的跟踪框发生漂移,且后者的跟踪框出现在视野之外;第 193 帧之后,目标发生了尺度变化、旋转和遮挡等,本文算法和 SRDCF 可以稳定地跟踪目标,且 SRDCF 对目标进行了更为准确的尺度估计,其他算法的跟踪情况仍然和第 193 帧时的跟踪情况相同,即无法确定目标的准确位置。

## 5 结 论

在相关滤波目标跟踪算法的目标函数中加入了显著感知自适应空间正则项和响应得分自适应时间正则项。利用显著感知自适应空间正则项,使空间正则权重在跟踪过程中与目标建立联系,在解决边界效应的同时获取可靠的滤波器惩罚系数。利用响应得分自适应时间正则项,使时间正则项在跟踪过程中自适应更新,降低了滤波器退化带来的影响。采用交替方向乘子法进行算法优化,分别求解出滤波器函数、空间正则权重和时间正则化参数。实验结果表明,所提算法的跟踪性能较优,且在复杂跟踪场景下具有良好的鲁棒性。今后可继续改进显著感知自适应空间正则项和响应得分自适应时间正则项,进一步提升算法的跟踪性能。

### 参 考 文 献

- [1] Lu H C, Li P X, Wang D. Visual object tracking: a survey [J]. *Pattern Recognition and Artificial Intelligence*, 2018, 31(1): 61-76.
- [2] Zhang W, Kang B S. Recent advances in correlation filter-based object tracking: a review [J]. *Journal of Image and Graphics*, 2017, 22(8): 1017-1033.  
张微, 康宝生. 相关滤波目标跟踪进展综述 [J]. *中国图象图形学报*, 2017, 22(8): 1017-1033.
- [3] 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.
- [4] 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-ECCV2012. Lecture Notes in Computer Science*. Berlin, Heidelberg: Springer, 2012, 7575: 702-715.
- [5] 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.
- [6] Danelljan M, Häger G, Shahbaz Khan F, et al. Accurate scale estimation for robust visual tracking [C] // *Proceedings of the British Machine Vision Conference 2014*, Nottingham. Durham, England, UK: British Machine Vision Association Press, 2014: 1-11.
- [7] Li Y, Zhu J K. A scale adaptive kernel correlation filter tracker with feature integration [C] // Agapito L, Bronstein M, Rother C. *Computer Vision-ECCV2014 Workshops. Lecture Notes in Computer Science*. Cham: Springer, 2015, 8926: 254-265.
- [8] Liu X Y, Wang Y M, Ma W N. Scale-adaptive correlation filter tracking algorithm based on FHOG and LBP features [J]. *Laser & Optoelectronics Progress*, 2020, 57(4): 041512.  
刘晓悦, 王云明, 马伟宁. 融合 FHOG 和 LBP 特征的尺度自适应相关滤波跟踪算法 [J]. *激光与光电子学进展*, 2020, 57(4): 041512.
- [9] Chen F L, Ding Q H, Chang Z, et al. Multi-scale kernel correlation filter algorithm for visual tracking based on the fusion of adaptive features [J]. *Acta Optica Sinica*, 2020, 40(3): 0315001.  
陈法领, 丁庆海, 常铮, 等. 自适应特征融合的多尺度核相关滤波目标跟踪 [J]. *光学学报*, 2020, 40(3): 0315001.
- [10] Zhu H J, Yu F Q. Feature-weight and scale adaptive algorithm for kernel correlation tracking [J]. *Laser & Optoelectronics Progress*, 2019, 56(10): 101501.

- 朱宏基, 于凤芹. 特征权值与尺度自适应的核相关跟踪算法[J]. 激光与光电子学进展, 2019, 56(10): 101501.
- [11] Danelljan M, Häger G, Khan F S, et al. Learning spatially regularized correlation filters for visual tracking[C]//2015 IEEE International Conference on Computer Vision (ICCV), December 7-13, 2015, Santiago, Chile. New York: IEEE Press, 2015: 4310-4318.
- [12] Dai K N, Wang D, Lu H C, et al. Visual tracking via adaptive spatially-regularized correlation filters [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2019: 4665-4674.
- [13] Li F, Tian C, Zuo W M, et al. Learning spatial-temporal regularized correlation filters for visual tracking [C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 4904-4913.
- [14] Qin Y, Lu H C, Xu Y Q, et al. Saliency detection via cellular automata[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 7-12, 2015, Boston, MA, USA. New York: IEEE Press, 2015: 110-119.
- [15] Hu Z H, Han Q, Li Q. Correlation filter tracking algorithm based on temporal awareness and adaptive spatial regularization[J]. Acta Optica Sinica, 2020, 40(3): 0315003.  
胡昭华, 韩庆, 李奇. 基于时间感知和自适应空间正则化的相关滤波跟踪算法[J]. 光学学报, 2020, 40(3): 0315003.
- [16] Wu Y, Lim J, Yang M H. Object tracking benchmark [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1834-1848.
- [17] Galoogahi H K, Fagg A, Lucey S. Learning background-aware correlation filters for visual tracking[C]//2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 1144-1152.
- [18] Hong S, You T, Kwak S, et al. Online tracking by learning discriminative saliency map with convolutional neural network[C]// ICML'15: Proceedings of the 32nd International Conference on International Conference on Machine Learning, July 6-11, 2015, Lille, France. Massachusetts: JMLR, 2015: 597-606.