

## 全局块与局部块协作的相关滤波目标跟踪算法

喻露<sup>1</sup> 胡剑锋<sup>1,2</sup> 姚磊岳<sup>1,2</sup>

1 南昌大学信息工程学院 南昌 330031

2 江西科技学院协同创新中心 南昌 330098

(lu\_yu0413@163.com)

**摘要** 针对传统相关滤波跟踪器在目标尺度变化和部分遮挡时效果不佳等问题,基于KCF提出了一种全局块与局部块协作的分块跟踪算法。该算法首先根据目标的外观特征,对跟踪目标进行水平或垂直分块,并分别训练两个局部滤波器和一个全局滤波器;然后在跟踪过程中使用局部滤波器对局部块进行跟踪,并根据局部块的跟踪结果对全局块的中心点位置进行初始预测。最后通过全局滤波器确定目标的最终位置,并将相关更新和尺度参数反馈给局部滤波器,以更新全局滤波器和局部滤波器。此外,不同于KCF使用单一的HOG特征,该算法合并了CN特征,改善了HOG特征对目标形变和运动模糊的表达能。另外,为解决目标部分遮挡导致的模型漂移问题,提出了一种基于有效局部块来指导模型更新的方法,并给出了有效局部块的评价标准。同时,该算法通过分析前后两帧局部块之间的距离变化对目标的尺度进行估计,解决了因目标尺度变化带来的跟踪失败问题。实验在包含100个视频序列的公共数据集OTB-100上进行,在评价指标上,以AUC得分为主,DP和OP为辅,对算法的性能进行评估。实验结果表明:所提出的算法能有效应对尺度变化和部分遮挡的问题,AUC得分在KCF的基础上提升了10%,总体性能也比KCF的其他4个改进算法更优,处理速度达到32fps。

**关键词:** 目标跟踪;相关滤波;分块;尺度变化;部分遮挡

中图法分类号 TP391

## Correlation Filter Object Tracking Algorithm Based on Global and Local Block Cooperation

YU Lu<sup>1</sup>, HU Jian-feng<sup>1,2</sup> and YAO Lei-yue<sup>1,2</sup>

1 School of Information Engineering, Nanchang University, Nanchang 330031, China

2 Center of Collaboration and Innovation, Jiangxi University of Technology, Nanchang 330098, China

**Abstract** Traditional correlation filter trackers are not effective in dealing with the problem that caused by target scale changing and partial occlusion. Aiming at sloving this problem, a block tracking algorithm based on KCF was proposed in this paper. In the first step, tracking object is divided horizontally or vertically according to its apperance feature. Then, in the tracking process, local filter is used to track local block, and center point position of the global block can be predicted by the tracked result of the local blocks. At last, the final position of the target is determined by the global filter. The relevant information renewal and scale parameters are fed back to the local filters to update both global and local filter. In addition, different from KCF, which only uses HOG feature, CN feature is imported in the proposed algorithm to enhance the ability of traget deformation tracing and motion blurring tracing. Moreover, in order to solve model drift problem caused by partial occlusion, a method based on effective local block is raised to guide model updating. Criteria of evaluating effective local block also defined. Furthermore, the scale of the target can be effectively estimated by analyzing the distance between local blocks, which solves tracking failure problem caused by target scale changing. The algorithm is evaluated on the public dataset OTB-100, which contains 100 video samples. The results show that the proposed algorithm performs quite well in the situation of scale changing and partial occlusion. Compared with KCF, the accuracy of the proposed algorithm is improved by 10%, and the overall performance is better than other four KCF

到稿日期:2019-05-17 返修日期:2019-09-09 本文已加入开放科学计划(OSID),请扫描上方二维码获取补充信息。

基金项目:国家自然科学基金(61762045);江西省科技厅项目(20171BAB202031);江西省科技厅科技攻关项目(20171BBE50060);江西省博士后援助项目(2017KY33);江西省教育厅项目(GJJ161143,GJJ151146);江西省科技厅科技计划专项重点研发项目(20181BBE50018)

This work was supported by the National Natural Science Foundation of China (61762045), Project of Department of Science and Technology of Jiangxi Province (20171BAB202031), Science and Technology Research Project of Jiangxi Science and Technology Agency (20171BBE50060), Postdoctoral Assistance Project of Jiangxi Province (2017KY33), Project of Department of Education of Jiangxi Province (GJJ161143, GJJ151146), and Science and Technology Plan Special Key Research and Development Project of Jiangxi Science and Technology Department (20181BBE50018).

通信作者:姚磊岳(lei\_yue\_yao@163.com)

based algorithms. The processing speed of the algorithm reaches 32 fps.

**Keywords** Visual tracking, Correlation filter, Block, Scale change, Partial occlusion

## 1 引言

目标跟踪的目的是在给定第一帧图像中目标的初始状态(通常为位置或范围)的情况下,对后续图像序列中每一帧的目标位置进行估计<sup>[1]</sup>。目标跟踪融合了包括计算机视觉、图像处理、信号处理、模式识别等众多领域的知识,在自动驾驶、智能监控、行为识别、人机交互等领域有着广泛的应用;并且随着硬件设施的不断完善和人工智能领域的快速发展<sup>[2]</sup>,其作为计算机视觉中的一个重要分支,已经成为了国内外研究人员重点研究对象。

相关滤波器(Correlation Filters, CF)在目标识别、图像配准、人脸识别和动作识别等计算机视觉应用中扮演着检测器的角色<sup>[1]</sup>。传统 CF 的设计目标是生成背景值低、场景感兴趣区域值高的响应图,在目标定位中效果显著;但因其训练要求不适用于目标跟踪,因此一直没有被应用于目标跟踪领域。Bolme 等于 2010 年提出的误差最小平方和滤波器(Minimum Output Sum of Squared Error, MOSSE)<sup>[3]</sup>,通过使用一种自适应的训练策略实时、有效地实现了目标跟踪。自此,CF 在目标跟踪领域大放异彩,成为了研究人员重点关注的对象。2012 年,Henriques 等在 MOSSE 的基础上提出了一种基于循环矩阵的核跟踪方法(Circulant Structure of Tracking with Kernels, CSK)<sup>[4]</sup>。CSK 从数学上完美解决了密集采样的问题,利用快速傅里叶变换(Fast Fourier Transform, FFT)实现了检测的过程。随后,Henriques 等又提出了基于多通道方向梯度直方图(Histogram of Oriented Gradient, HOG)<sup>[5]</sup>特征的核相关滤波器(Kernelized Correlation Filters, KCF)<sup>[6]</sup>。同年,Danelljan 等也提出了将自适应颜色特征加入 CSK 的 CN (Color Name)<sup>[7]</sup>算法。

上述传统 CF 类算法都没有考虑到尺度变化和遮挡的问题,而尺度变化是跟踪中比较基本和常见的问题,无论目标的尺度是变小还是变大,如果不能及时应对,都会导致模型漂移和跟踪失败。基于此,研究人员提出了使用尺度池技术来应对尺度变化的问题,其中的代表作有多特征尺度自适应跟踪器(Scale Adaptive with Multiple Features tracker, SAMF)<sup>[8]</sup>和判别尺度空间跟踪器(Discriminative Scale Space Tracking, DSST)<sup>[9]</sup>。SAMF 是在 KCF 的基础上进行的改进,采用多特征(灰度、HOG、CN)融合,并且在跟踪定位的过程中使用了类似金字塔的多尺度搜索策略。此后,该策略被广泛应用于 CF 类跟踪算法<sup>[10-13]</sup>。而 DSST 采用了与 SAMF 完全不同的策略来应对尺度变化,该算法是把目标跟踪看成目标中心平移和目标尺度变化两个独立问题,同时训练两个滤波器,两个滤波器相互依赖,从而很好地缓解了尺度变化带来的精度下降等问题;并且其作者在文献<sup>[14-19]</sup>中对算法不断进行改进,使得此方法被大量应用于 CF 类跟踪器<sup>[20-21]</sup>。与此同时,遮挡也是跟踪领域里亟需解决的问题;当目标发生遮挡时,如果依旧更新模型,将会使得模型被污染,导致模型漂移,后续帧无法跟踪目标。为此,出现了大量的遮挡评价标准,

MOSSE 使用了峰值旁瓣比(Peak to Sidelobe Ratio, PSR)<sup>[3]</sup>;基于熵最小化的多专家鲁棒(Multiple Experts using Entropy Minimization, MEEM)<sup>[22]</sup>跟踪器通过多专家修复方案防止模型漂移;基于循环特征的大边际目标跟踪(Large Margin Object Tracking with Circulant Feature Maps, LMCF)<sup>[23]</sup>算法和文献<sup>[24]</sup>使用了平均峰值相关能量(Average Peak-to Correlation Energy, APCE)衡量跟踪质量;长期的相关跟踪算法(Long-term Correlation Tracking, LCT)<sup>[25]</sup>在模型中加入了检测目标置信度的滤波器;具有信道和空间可靠性的判别相关滤波器(Discriminative Correlation Filter with Channel and Spatial Reliability, CSR-DCF)<sup>[26]</sup>将次峰值和主峰值之比作为检测跟丢的衡量标准。

尺度搜索策略和各种应对遮挡的评价指标在一定程度上促进了相关滤波类算法的发展,但它们也有其局限性。尺度池技术严重依赖于尺度因子的个数,个数越多,算法的精度也就越高,但同时也带来了巨大的时间消耗,使得相关滤波类算法的天然优势大大减弱。而各种评价指标也只有针对特定方法才能够带来较好的效果。

针对上述问题,本文在 KCF 的基础上提出了全局块与局部块协作的相关滤波器(Cooperation between Global and Local Correlation Filter, CGLCF),算法模型如图 1 所示。该模型依赖于全局滤波器和两个局部滤波器之间的协作交互作用,局部滤波器为目标最终位置提供一个初始估计,再通过全局滤波器确定最终结果;并且根据有效局部块来指导模型的更新,防止模型漂移;另外,跟踪器利用前后两帧局部块之间的距离变化估计目标的尺度信息。因此,该模型不仅可以应对部分遮挡,还可以处理尺度变化,有较好的鲁棒性。

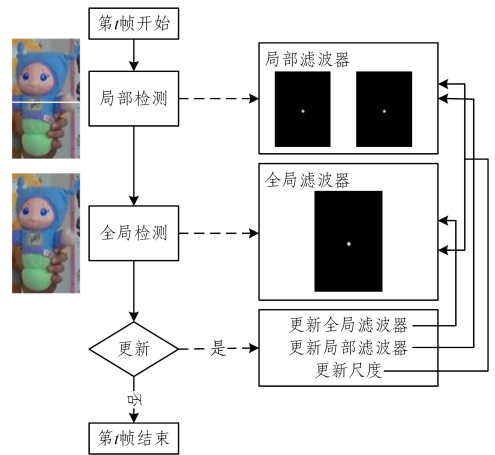


图 1 CGLCF 的模型

Fig. 1 Model of CGLCF

## 2 全局块与局部块协作的相关滤波器算法

CGLCF 算法针对传统相关滤波器(如 MOSSE, CSK, KCF 等)在目标尺度变化和部分遮挡时性能不佳等问题,在 KCF 的基础上采用分块跟踪的策略,提出通过前后两帧局部

块之间的距离变化初步估计目标当前帧的尺度因子,再将其与初始化以来所有帧的尺度因子的平均值进行比较,以确定是否对目标的最终尺度进行更新;并且使用有效局部块对模型进行自适应更新,防止模型漂移;同时合并了 CN 特征,改善了 HOG 特征对目标形变和运动模糊的表达能力。

### 2.1 KCF

在 KCF 中, Henriques 等将相关滤波器的求解转化为岭回归问题,采用循环矩阵对样本进行密集采样,引入核方法来提升相关滤波算法的鲁棒性,使其能够处理非线性分类问题<sup>[1]</sup>。

KCF 跟踪器在  $M \times N$  像素的图像块  $x$  上训练相关滤波器,其中  $x_{m,n}$  为  $x$  经过循环移位生成的训练样本,并且样本标签使用高斯函数  $y_{m,n}(m,n) \in \{0,1,\dots,M-1\} \times \{0,1,\dots,N-1\}$  表示,然后通过式(1)计算得到  $w$ 。

$$w = \arg \min_w \sum_{m,n} |\langle \phi(x_{m,n}), w \rangle - y_{m,n}|^2 + \lambda \|w\|^2 \quad (1)$$

其中,  $\phi$  为非线性映射函数,  $\lambda$  为正正则化参数。

利用 FFT 将式(1)最小化为  $w = \sum_{m,n} a(m,n) \phi(x_{m,n})$ , 其中系数  $a$  采用式(2)计算。

$$a = F^{-1} \left( \frac{F(y)}{F(\langle \phi(x), \phi(x) \rangle) + \lambda} \right) \quad (2)$$

其中,  $y = \{y_{m,n} | (m,n) \in \{0,1,\dots,M-1\} \times \{0,1,\dots,N-1\}\}$ ,  $F$  和  $F^{-1}$  分别表示傅里叶变换及其逆变换。根据所学习到的  $a$  和外观模型  $\hat{x}$ , 可以通过式(3)计算出当前帧的图像块  $z$  的响应图。

$$\hat{y} = F^{-1}(F(a) \odot F(\langle \phi(z), \phi(\hat{x}) \rangle)) \quad (3)$$

其中,  $\odot$  表示矩阵乘积, 目标位置由  $\hat{y}$  中最大值所在位置决定。

### 2.2 全局块与局部块的协作模型

对于全局块,采用 KCF 中的全局外观模型。对于局部块,本文使用了一种简单有效的自适应分块策略<sup>[27]</sup>:若目标的长大于宽,则将目标沿垂直方向二等分(如图 2(a)所示);否则,将目标沿水平方向二等分(如图 2(b)所示)。



(a)垂直分块

(b)水平分块

图 2 分块策略示例图

Fig. 2 Example diagram of blocking strategy

对跟踪目标进行自适应分块后,对初始帧中的全局块和局部块分别训练一个全局滤波器和两个局部滤波器。跟踪时,滤波器通过全局块与局部块的联合作用预测目标的最终位置,这种协作模型有助于对目标对象进行鲁棒跟踪。

具体来说,首先,用每个局部滤波器检测对应的局部块,再基于局部块的检测结果对目标对象(即全局块)的中心位置进行初始预测。接着,使用这个初始预测定义全局块搜索区域,并通过全局滤波器来确定目标的最终位置。最后,根据有

效局部块指导模型的更新,并通过前后两帧局部块之间的距离变化估计目标的尺度信息。可以看出,全局滤波器和局部滤波器在模型中发挥着同等重要的作用。

### 2.3 模型更新

模型更新是所有跟踪器需要解决的一个难点。在目标跟踪过程中,由于光照、姿态等因素的影响,目标的外观会发生变化。因此,跟踪器需要更新模型,以适应外观变化。然而,对于相关滤波类跟踪器,每次检测的结果都有一定误差,如果模型更新过于频繁,这个误差会随着跟踪时间慢慢积累,由量变到质变,最终在后续帧丢失目标。在 KCF 跟踪器中,模型由学习到的目标外观  $x$  和分类器系数  $\alpha$  组成,并采用简单的线性插值方式来更新  $x$  和  $\alpha$ :  $x' = (1 - \gamma)x^{t-1} + \gamma x$ ,  $\alpha' = (1 - \gamma)\alpha^{t-1} + \gamma\alpha$ 。由于 KCF 使用固定的学习率  $\gamma$ ,因此  $x$  和  $\alpha$  将在不适应特定序列帧的情况下进行持续更新。如果跟踪器失去目标,整个模型将在剩余帧中受到污染,从而将目标跟丢。CGLCF 将使用下文的自适应更新方案来处理这个问题。

对于全局块,在模型更新前,首先对局部块  $i$  的有效性进行检测,将检测分数与阈值进行比较,如检测分数大于阈值,则判断该局部块为有效块。如式(4)所示,如果存在有效局部块,则对模型进行更新;否则,保留前一帧的模型,不进行更新。

$$\alpha^t = \begin{cases} (1 - \gamma)\alpha^{t-1} + \gamma\alpha, & \text{存在有效局部块} \\ \alpha^{t-1}, & \text{无有效局部块} \end{cases} \quad (4)$$

$$x^t = \begin{cases} (1 - \gamma)x^{t-1} + \gamma x, & \text{存在有效局部块} \\ x^{t-1}, & \text{无有效局部块} \end{cases}$$

对于局部块,同样地,在模型更新前,对局部块的有效性进行检测。如果该局部块为有效块,则对该局部块的模型进行更新;否则,保留之前的模型,不进行更新。

$$\alpha_i^t = \begin{cases} (1 - \gamma)\alpha_i^{t-1} + \gamma\alpha_i, & \text{局部块 } i \text{ 有效} \\ \alpha_i^{t-1}, & \text{局部块 } i \text{ 无效} \end{cases} \quad (5)$$

$$x_i^t = \begin{cases} (1 - \gamma)x_i^{t-1} + \gamma x_i, & \text{局部块 } i \text{ 有效} \\ x_i^{t-1}, & \text{局部块 } i \text{ 无效} \end{cases}$$

同时,本文提出使用最大置信度均值比(Mean Ratio of Max Confidence, MRMC),即目标在第  $t$  帧中的最大置信度值与第 1 至  $t-1$  帧中最大置信度的平均值的比值,来对局部块的有效与否进行判定,计算方法如式(6)所示。

$$\text{MRMC} = \frac{\max(\hat{f}(z_i^t))}{\frac{1}{t-1} \sum_{n=1}^{t-1} \max(\hat{f}(z_i^n))} \quad (6)$$

其中,  $\hat{f}(z_i^t)$  为第  $t$  帧中局部块  $i$  的所有候选目标子块的响应值,  $\max(\hat{f}(z_i^t))$  为它的最大响应值,  $\max(\hat{f}(z_i^t))$  所在的位置即为局部块  $i$  在第  $t$  帧的跟踪结果。一般情况下,当目标发生遮挡时,  $\max(\hat{f}(z_i^t))$  会变小。因此,用 MRMC 判断局部块是否发生遮挡,如 MRMC 小于阈值  $\phi$ , 则判定该局部块出现遮挡,为无效局部块。

通过上述模型更新方式,本文算法在目标发生遮挡或外观发生变化的情况下更加鲁棒,即使局部甚至全局对象被遮挡,也能够准确地跟踪目标。

## 2.4 尺度变化

使用自适应的方式改变滤波器的大小,以应对跟踪目标不同帧不同尺度的问题,具有重要意义。大多数基于相关滤波器的跟踪器都试图通过使用一个判别滤波器和一个基于金字塔结构的尺度搜索池来估计目标对象的大小。尽管这些跟踪器取得了显著的结果,但它们无法估计当前帧中目标对象的大小。

本文提出使用基于分块的模式自适应地改变滤波器的大小,以便进行更准确的检测。众所周知,前后帧中两个局部块之间的距离变远,说明跟踪目标的尺度在变大;类似地,前后帧中两个局部块之间的距离变近,说明目标的尺度在缩小。该尺度估计方案具有鲁棒性,它提供了一种简单有效的估计目标尺度的方法。

在不考虑以前尺度信息的情况下,每一帧都改变目标尺度可能会导致估计不准确。因此,只有当尺度稳定地增加或减少时,模型才会用尺度因子  $s$  来改变原始滤波器的大小;并且一旦尺度进行更新,模型将根据估计的尺度因子  $s$  更改滤波器大小来重新初始化模型。使用式(7)计算第  $t$  帧的尺度因子:

$$s^t = \frac{\| \ell_i^t - \ell_j^t \|}{\| \ell_i^{t-1} - \ell_j^{t-1} \|} \quad (7)$$

其中,  $i$  和  $j$  表示局部块,  $\ell_i^t$  和  $\ell_j^t$  分别为局部块  $i$  和  $j$  在第  $t$  帧的位置,  $\| \ell_i^t - \ell_j^t \|$  为  $\ell_i^t$  和  $\ell_j^t$  之间的距离,  $\| \ell_i^{t-1} - \ell_j^{t-1} \|$  为第  $t-1$  帧中局部块之间的距离。同时,所提出的算法模型中存储了重新初始化以来所有序列帧的尺度因子,并提出使用重新初始化以来所有序列帧的尺度因子的平均值  $\eta = \text{avg}(s^{0:t})$ ,来判断是否更新当前帧目标尺度和滤波器的大小。如果  $\eta < 1 + \delta$  或者  $\eta > 1 - \delta$ ,并且局部块都有效,则表示需要更新目标尺度和滤波器的大小。实验证明,这种简单的尺度策略在实际应用中是有效的。

## 2.5 算法流程

本文算法的具体执行步骤如算法 1 所示。

### 算法 1 CGLCF

Step1 根据目标的外观特征进行垂直或水平分块。

Step2 局部块目标检测。

Step2.1 用  $c_i^{t-1}$  表示第  $t-1$  帧中第  $i$  个局部块位置的中心点,在第  $t$  帧中提取以  $c_i^{t-1}$  为中心、大小为局部块 2.4 倍的图像块作为搜索区域,计算其 HOG+CN<sup>[7]</sup> 特征  $z_i^t$ 。

Step2.2 根据式(3)计算第  $i$  个局部块的所有候选样本的响应值  $\hat{f}(z_i^t)$ 。

Step2.3 找到  $\hat{f}(z_i^t)$  中最大值  $\hat{f}_{\max}(z_i^t)$  对应的偏移量  $\Delta c = (\Delta c_h, \Delta c_v)$ ,得到第  $t$  帧中第  $i$  个局部块的跟踪结果  $c_i^t = c_i^{t-1} + \Delta c$ 。

Step3 全局块目标检测。

Step3.1 在第  $t$  帧中,根据第  $i$  个局部块的跟踪结果  $c_i^t = c_i^{t-1} + \Delta c$ ,计算局部块之间的中心点,得到全局块初始估计的中心点  $c^t$ ,并取大小为全局块 2.4 倍的图像块作为搜索区域,计算其 HOG+CN 特征  $z^t$ 。

Step3.2 根据式(3)计算全局块的所有候选目标位置的响应值  $\hat{f}(z^t)$ 。

Step3.3 找到  $\hat{f}(z^t)$  中最大值  $\hat{f}_{\max}(z^t)$  对应的偏移量  $\Delta g = (\Delta g_x, \Delta g_v)$ ,得到第  $t$  帧中的跟踪结果  $g^t = c^t + \Delta g$ 。

Step4 更新尺度。根据式(7)计算第  $t$  帧的尺度因子并保存,计算尺度因子的平均值  $\eta = \text{avg}(s^{0:t})$ ,使用第 2.4 节的方法改变全局块尺度,并更新局部块的尺度信息。

Step5 局部块分类器的训练。

Step5.1 在第  $t$  帧中提取以  $c_i^t$  为中心、大小为局部块 2.4 倍的图像块作为搜索区域,计算其 HOG+CN 特征  $x_i^t$ 。

Step5.2 根据式(2)计算  $a_i$ 。

Step6 局部块模型更新。根据式(6)计算第  $i$  个局部块的 MRMC,并根据式(5)计算  $\alpha_i^t$  和  $x_i^t$ 。

Step7 全局分类器的训练。

Step7.1 在第  $t$  帧中提取以  $g^t$  为中心、大小为全局块 2.4 倍的图像块作为搜索区域,计算其 HOG+CN 特征  $x^t$ 。

Step7.2 根据式(2)计算  $a$ 。

Step8 全局块模型的更新。根据式(5)计算  $\alpha^t$  和  $x^t$ 。

Step9 重复 Step2-Step8,直至最后一帧。

## 3 实验

### 3.1 实验平台与参数配置

实验平台:Windows 10+Matlab 2016b;计算机配置:Intel(R) Core(TM) i7-7700HQ,2.80GHz CPU,8GB RAM。

实验的参数与 KCF 中的参数基本一致。由于 HOG 特征在低分辨率(30×30 像素以下)下表现不佳,因此本文实验中,在分辨率低于 30 的图像上,设置 cell 的大小为 2×2。图像搜索区域设置为跟踪目标的 2.4 倍。

理论上,局部块的有效性数值在 0.5 到 1 之间;而通过实验观察到有效性数值在 0.5 到 0.75 之间,并且发现 MRMC 在 0.56 左右时出现严重遮挡、快速运动、变形等情况。因此,我们将第 2.3 节的阈值  $\phi$  设置为 0.55。另外,与使用尺度池技术的算法不同,在本文算法中,如果尺度发生变化,滤波器的大小也将改变。这意味着,尺度大小更改得越多,模型就越容易忘记前面的信息。因此,我们将第 2.4 节的尺度阈值  $\delta$  设置为 0.15。

### 3.2 评价指标

为了评估本文算法的鲁棒性和准确性,在公共基准数据集 OTB-100<sup>[28]</sup> 上进行实验。该数据集包含 100 个视频数据,其中囊括了 11 个不同的挑战,如光照变化(Illumination Variation, IV)、尺度变化(Scale Variation, SV)、遮挡(Occlusion, OCC)、变形(Deformation, DEF)、运动模糊(Motion Blur, MB)、快速运动(Fast Motion, FM)、平面内旋转(In-Plane Rotation, IPR)、平面外旋转(Out-of-Plane Rotation, OPR)、视野外(Out-of-View, OV)、背景杂乱(Background Clutters, BC)和低分辨率(Low Resolution, LR)。在实验中,采用中心位置误差(Center Location Error, CLE)(也称距离精度(Distance Precision, DP))和重叠率(Overlap Precision, OP)两种评价标准对算法性能进行测试。

中心位置误差(CLE):被跟踪目标的中心位置与人工标注的中心位置之间的平均欧氏距离。通常,用一个视频序列所有帧的平均中心位置误差来评定该序列的总体性能。基于 CLE,文献[28]提出了便于直观理解的精确度图(Precision

Plot)。精确度图是由一个视频序列中跟踪器定位的位置与真实人工标注的位置的差值小于某个阈值的帧数比例所汇总绘制而成的,通常采用阈值为 20 像素内达到要求的帧数比例的得分来衡量跟踪器的整体性能。但是,当跟踪器丢失目标时,跟踪输出的位置可能是随机的,因此平均误差值可能不能正确测试跟踪性能。

重叠率(OP):大多数研究人员将 OP 作为跟踪器的主要评价指标。给定被跟踪的边界框  $r_t$  和人工标注真实边界框  $r_a$ ,重叠得分定义为  $s = \frac{|r_t \cap r_a|}{|r_t \cup r_a|}$ ,其中  $\cap$  和  $\cup$  分别表示两个区域的交集和并集,  $|\cdot|$  表示区域内的像素个数。为了测试序列帧的性能,计算重叠率大于给定阈值的成功帧数,阈值  $t_o$  越大,说明对重叠率  $s$  的要求越高,因此重叠率随着阈值的增大而下降。基于 OP,文献[28]又提出了成功率图(Success Plot)。成功率图显示了阈值为 0 到 1 之间的成功帧的比率。然而,在跟踪器评估的特定阈值处使用一个成功率值可能不公平或不具有代表性。因此,本文使用成功率图中曲线下的面积(Area Under Curve,AUC)来对跟踪算法的性能进行评估。

### 3.3 实验结果

为了评估算法的性能,基于 OTB-100 数据集,从整体性

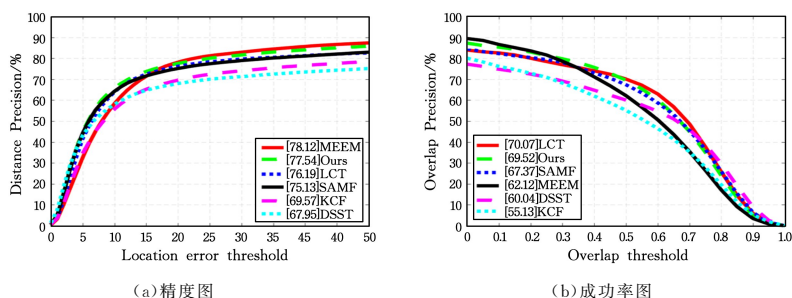


图 3 6 个跟踪器的 DP 和 OP 得分对比

Fig. 3 Comparison of DP and OP scores of 6 trackers

表 1 6 个跟踪器的 AUC 得分对比

Table 1 Comparison of AUC scores of 6 trackers

(单位:%)

跟踪器	All	IV	SV	OCC	DEF	MB	FM	IPR	OPR	OV	BC	LR
Ours	<b>56.45</b>	<b>58.51</b>	<b>51.62</b>	<b>54.46</b>	<b>52.06</b>	<b>57.29</b>	<b>53.28</b>	<b>52.44</b>	<b>54.48</b>	<b>48.20</b>	<b>57.47</b>	<b>41.71</b>
KCF	47.71	48.23	39.47	44.47	43.79	45.86	45.90	46.94	45.31	39.33	49.77	30.74
DSST	51.27	55.62	46.56	44.90	41.44	46.88	44.68	50.22	46.98	38.59	52.27	38.33
SAMF	55.32	53.04	<u>49.21</u>	<u>53.75</u>	<u>50.48</u>	52.55	50.66	51.87	53.63	48.03	52.48	<b>42.99</b>
MEEM	52.96	52.18	47.21	50.75	49.22	<u>55.64</u>	<b>54.22</b>	<u>52.89</u>	52.54	<b>48.79</b>	51.89	36.41
LCT	<u>56.16</u>	<u>56.68</u>	48.63	50.59	49.71	53.31	<u>53.45</u>	<b>55.70</b>	<u>53.82</u>	45.17	<u>54.95</u>	35.43

#### 3.3.2 遮挡

在数据集 OTB-100 中,有 48 个视频序列存在不同程度的遮挡。图 4(a)~图 4(c)显示了 6 种算法在 Coke, Walking2, Girl2 这 3 个存在遮挡情况的视频序列上的实验结果。按遮挡程度的不同,可以将 3 个视频进行如下划分:Coke 为轻微遮挡,Walking2 为中度遮挡,Girl2 为严重遮挡。如图 4 所示,本文算法由于使用了新的模型更新方案,因此可以较灵活地应对不同程度的遮挡。在发生轻微遮挡时,由于相关滤波类算法扩大了检测区域,因此大多数跟踪器依然能跟上目标。但是,KCF,SAMF,DSST,LCT 等传统相关滤波类方法由于没有加入遮挡检测机制,在发生严重遮挡时仍然会更新

能、遮挡、尺度变化 3 个方面将本文算法与 5 个基于相关滤波的跟踪器(KCF,DSST,SAMF,MEEM,LCT)进行比较。

#### 3.3.1 整体性能

在 OTB-100 上,6 个跟踪器的 DP 和 OP 得分对比结果如图 3 所示。表 1 列出了 6 种算法重叠成功率图中 AUC 的得分,包含了整体性能与 11 个属性子集的性能,其中加粗的字体为最优性能,下划线字体为次优。从图 3 和表 1 中可以看出,虽然本文算法在 DP 和 OP 上不是最佳,均排在第二,但是 AUC 的整体性能值为 56.45,在 6 个跟踪器中最优。本文算法的 AUC 值在 IV,SV,OCC,DEF,MB,OPR,BC 属性子集中均为最大,在 OV 与 LR 中为第二,只有 FM 和 IPR 的 AUC 值不在前二,表明本文算法能够较好地应对各种突发情况,有较好的鲁棒性。同时,在尺度发生变化时,本文算法的 AUC 得分为 51.62,比基准算法 KCF 高出近 12 个百分点;而在发生遮挡时,本文算法的 AUC 值为 54.46,较 KCF 也高出了 10 个百分点,比使用尺度池技术的 DSST 和 SAMF 高出 5 个百分点以上。因此,本文所提出的模型在应对尺度变化和遮挡时有较好的效果;除此之外,本文算法的平均跟踪速度为 32 帧/s,能够达到实时性要求。

模型,导致模型被污染,从而使滤波器学习到错误的信息,进而无法跟踪到目标,如图 4(c)所示。而本文所提出的跟踪器和 MEEM 跟踪器在模型中加入了遮挡检测机制,对模型的更新方案进行了改进,因此即使在目标发生严重遮挡甚至完全遮挡时,依旧能够在后期跟踪到目标,不会发生严重漂移等问题。图 5(a)为 OTB-100 中遮挡属性视频子集的重叠成功率,可以看出,本文所提出的 CGLCF 跟踪器在发生遮挡时的性能高于其他 5 个跟踪器。

#### 3.3.3 尺度变化

在数据集 OTB-100 中,有 63 个视频序列的目标存在不同程度的尺度变化。图 4(b)、图 4(d)~图 4(f)给出了 6 种算

法在 Walking2, Singer1, Car4, Dog1 这 4 个存在明显尺度变化的视频序列上的实验结果。从图中可以看出, KCF 等一系列跟踪器由于没有使用尺度估计的策略, 跟踪框的大小从始至终都是一样的, 当目标的尺度变小或扩大时, 都很可能导致跟踪的漂移和失败; 而 CGLCF, SAMF 和 DSST 加入了多尺度检测方法, 因此当目标扩大或缩小时, 跟踪框会相应地扩大或缩小。另外, 图 4(f) 显示, 当目标的尺度变化较大时, SAMF 跟踪器无法准确、快速地对跟踪框进行改变, 这是由于 SAMF 仅使用了步长为 1.005 的 7 个尺度因子, 而如果增加尺度因子的数量, 该算法的实时性将无法保证。DSST 将目标跟踪看成目标中心平移和目标尺度变化两个独立问题, 单独训练了一个滤波器来对目标的尺度变化进行检测, 因而可以较鲁棒地应对目标的尺度变化。但是 DSST 对尺度因子的个数极其敏感, DSST 中标配的尺度因子为 33 个, 如果减少尺度因子数量, 滤波器将跟不上目标的尺度变化。

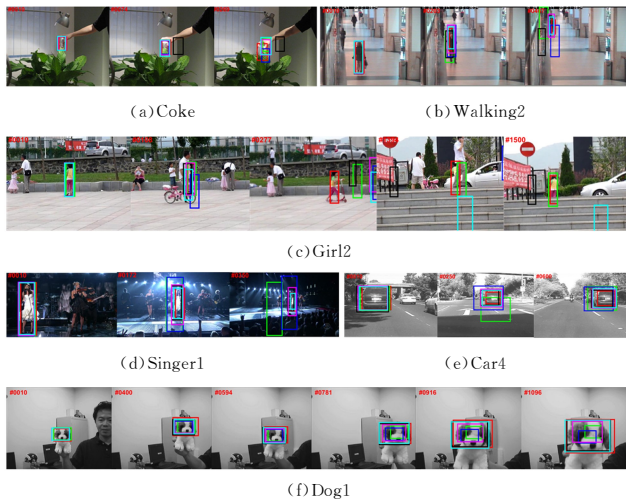


图 4 6 个算法的跟踪实例图

Fig. 4 Tracking example diagram of 6 algorithms

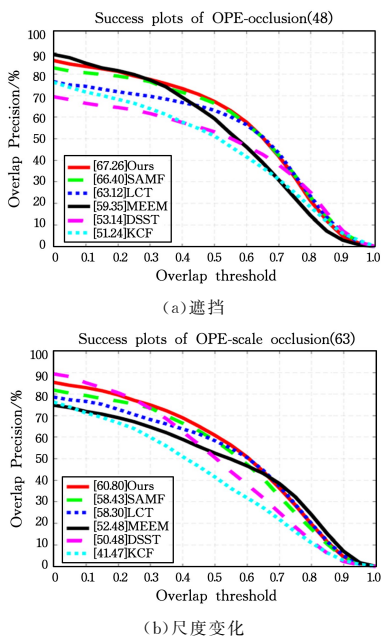


图 5 6 种算法在遮挡和尺度变化属性视频子集的成功率图

Fig. 5 Comparison of success rates of six algorithms in occlusion and scale change attribute video subsets

从图 5(b)中可以看出, 本文所提出的使用局部块位置的中心距离来判断是否对尺度进行变化, 并对滤波器大小进行更新的方案, 能够较为快速、准确地把握目标的尺度变化, 在 63 个发生尺度变化的视频序列上性能最优。

**结束语** 本文在 KCF 的基础上, 提出将目标对象分为全局块与局部块, 并同时训练 3 个滤波器, 接着对局部块进行跟踪, 将局部块的跟踪结果作为全局块的最终位置的参考值, 最后依据全局滤波器得到最终的跟踪结果。由于使用了有效局部块检测和新的模型更新方案, 并依据前后两帧局部块之间的距离变化估计目标尺度的变化, 本文提出的 CGLCF 算法在目标被部分遮挡和尺度变化时有较好的鲁棒性。实验证明, CGLCF 算法的精度在基准算法 KCF 的基础上提升了 10 多个百分点, 并且比使用尺度池技术的跟踪器更加简单有效。但是, 该算法也存在不足: 由于训练了 3 个滤波器, 其在时间效率上不如 KCF。后续工作将从时间效率的角度改进算法, 提升算法的处理速度。

## 参考文献

- [1] 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.
- [2] MENG L, YANG X. A Survey of Object Tracking Algorithms [J]. Acta Automatica Sinica, 2019, 45(7): 1244-1260.
- [3] BOLME D S, BEVERIDGE J R, DRAPER B A, et al. Visual Object Tracking using Adaptive Correlation Filters [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2010: 2544-2550.
- [4] HENRIQUES J F, CASEIRO R, MARTINS P, et al. Exploiting the Circulant Structure of Tracking-by-Detection with Kernels [C] // Proceedings of the 12th European Conference on Computer Vision. Heidelberg: Springer, 2012, 7575: 702-715.
- [5] DALAL N, TRIGGS B. Histograms of Oriented Gradients for Human Detection [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2005: 886-893.
- [6] 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.
- [7] DANELLJAN M, KHAN F S, FELSBERG M, et al. Adaptive Color Attributes for Real-time Visual Tracking [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2014: 1090-1097.
- [8] LI Y, ZHU J K. A Scale Adaptive Kernel Correlation Filter Tracker with Feature Integration [C] // Proceedings of the European Conference on Computer Vision. Heidelberg: Springer, 2014: 254-265.
- [9] DANELLJAN M, HAGER G, KHAN F S, et al. Discriminative Scale Space Tracking [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(8): 1561-1575.

- [10] MA C, HUANG J B, YANG X K, et al. Hierarchical Convolutional Features for Visual Tracking[C] // Proceedings of the IEEE International Conference on Computer Vision. Los Alamitos; IEEE Computer Society Press, 2015:3074-3082.
- [11] BERTINETTO L, VALMADRE J, GOLODETZ S, et al. Staple; Complementary Learners for Real-time Tracking[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos; IEEE Computer Society Press, 2016:1401-1409.
- [12] GALOOGAHI H K, FAGG A, LUCEY S. Learning Background-Aware Correlation Filters for Visual Tracking[C] // Proceedings of the IEEE International Conference on Computer Vision. Los Alamitos; IEEE Computer Society Press, 2017:1144-1152.
- [13] MUELLER M, SMITH N, Ghanem B. Context-Aware Correlation Filter Tracking[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos; IEEE Computer Society Press, 2017:1387-1395.
- [14] DANELLJAN M, HAGER G, KHAN F S, et al. Learning Spatially Regularized Correlation Filters for Visual Tracking[C] // Proceedings of the IEEE International Conference on Computer Vision. Los Alamitos; IEEE Computer Society Press, 2015:4310-4318.
- [15] DANELLJAN M, HAGER G, KHAN F S, et al. Accurate Scale Estimation for Robust Visual Tracking[C] // Proceedings of British Machine Vision Conference. Swansea; BMVA Press, 2014:65. 1-65. 11.
- [16] DANELLJAN M, HAGER G, KHAN F S, et al. Adaptive Decontamination of the Training Set; A Unified Formulation for Discriminative Visual Tracking[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos; IEEE Computer Society Press, 2016:1430-1438.
- [17] DANELLJAN M, BHAT G, KHAN F S, et al. ECO; Efficient Convolution Operators for Tracking[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos; IEEE Computer Society Press, 2017:6931-6939.
- [18] DANELLJAN M, HAGER G, KHAN F S, et al. Convolutional Features for Correlation Filter based Visual Tracking[C] // Proceedings of IEEE International Conference on Computer Vision Workshop. Los Alamitos; IEEE Computer Society Press, 2015:621-629.
- [19] BHAT G, JOHNANDER J, DANELLJAN M, et al. Unveiling the Power of Deep Tracking[C] // Proceedings of the European Conference on Computer Vision. Heidelberg; Springer, 2018:493-509.
- [20] ZHU H J, ZHU F, FAN Z F, et al. Visual Object Tracking Method with Motion Estimation and Scale Estimation[J]. Computer Science, 2017, 44(S2):193-198.
- [21] QU Z, ZHAO C M. Anti-occlusion Adaptive-scale Object Tracking Algorithm [J]. Computer Science, 2018, 45(4):296-300.
- [22] ZHANG J M, MA S G, SCLAROFF S. MEEM; Robust Tracking via Multiple Experts using Entropy Minimization[M] // Lecture Notes in Computer Science. Heidelberg; Springer, 2014, 8694:188-203.
- [23] WANG M M, LIU Y, HUANG Z Y. Large Margin Object Tracking With Circulant Feature Maps[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos; IEEE Computer Society Press, 2017:4800-4808.
- [24] JING X S, ZOU W J, XIA T, et al. Adaptive Complementary Learners with Diversified Color Attributes for Object Tracking [J]. Journal of Computer-Aided Design & Computer Graphics, 2018, 30(12):63-71.
- [25] MA C, YANG X K, ZHANG C Y, et al. Long-Term Correlation Tracking[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos; IEEE Computer Society Press, 2015:5388-5396.
- [26] LUKEZIC A, VOJIR T, ZAJC L C, et al. Discriminative Correlation Filter with Channel and Spatial Reliability[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos; IEEE Computer Society Press, 2016:4847-4856.
- [27] MAO N, YANG D D, CAI Y Z, et al. Reliable Patch Tracking Based on Multi-feature Correlation Filter[J]. Control Engineering of China, 2018, 25(4):650-656.
- [28] WU Y, LIM, YANG M H. Object Tracking Benchmark [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9):1834-1848.



**YU Lu**, born in 1994, postgraduate. His main research interests include computer vision, machine learning and object tracking.



**YAO Lei-yue**, born in 1982, Ph.D, professor, postgraduate supervisor. His main research interests include computer vision and information processing.