

文章编号:1007-2780(2019)11-1115-09

连续域卷积操作跟踪算法的改善

王 鑫^{1*}, 李玉芳², 宋 策¹, 韩松伟¹

(1. 中国科学院长春光学精密机械与物理研究所, 吉林 长春 130033;

2. 中国第一汽车集团有限公司研发总院, 吉林 长春 130013)

摘要:针对连续域卷积操作跟踪算法中存在的计算复杂问题和过拟合问题,提出一种降维卷积因式分解方法和一个紧凑衍生模型。本文选取连续域卷积操作跟踪算法中的部分特征通道、因式分解卷积算子,并结合高斯-牛顿法和共轭梯度迭代求取降维矩阵和分类器,以达到降低系统的计算复杂程度的目的。选取服从高斯分布的若干样本,结合联合概率密度函数及混合高斯分布模型,构建一个更具代表性的紧凑衍生模型,既能降低过拟合出现的风险又能保证样本的多样性。通过降低模型更新频率,减少计算复杂度,同时降低模型漂移发生的概率。经实验验证,本文基于四类标准数据集获得的跟踪性能曲线积分比连续域卷积操作跟踪算法最少高出 0.8%,最多高出 13%。结果表明,降维卷积因式分解法和紧凑衍生模型可以有效地提升系统跟踪性能。

关键词:连续域卷积操作;卷积因式分解;过拟合;紧凑衍生模型

中图分类号:TP391.4 **文献标识码:**A **doi:**10.3788/YJYXS20193411.1115

Algorithm improvement strategy about continuous convolution operators for visual tracking

WANG Xin^{1*}, LI Yu-fang², SONG Ce¹, HAN Song-wei¹(1. *Changchun Institute of Optics, Fine Mechanics and Physics,*
*Chinese Academy of Science China, Changchun 130033, China;*2. *Research and Development Institute, China Faw Group Corporation,*
Changchun 130013, China)

Abstract: The reduced dimensional factorized convolution method and the compact generative model (RFCG) are proposed in order to solve the computational complexity and over-fitting problems in the continuous convolution operators for visual tracking(C-COT). Some feature channels in the C-COT are selected in advance, then the convolution operator is factorized, and the reduced dimension matrix and filter are obtained by combining Gauss-Newton method and conjugate gradient iteration (CG), so as to reduce the computational complexity of the system in this paper. A number of samples subject to the Gaussian distribution are selected and combined with the joint probability density function and the mixed Gaussian distribution model to build a compact generative model with stronger representativeness. The risk of the over-fitting problems can be reduced and the diversity of samples can be guaran-

收稿日期:2019-04-09;修订日期:2019-09-11.

* 通信联系人, E-mail:14S104020@hit.edu.cn

teed in this way. Finally, the model updating frequency is reduced to save the calculation time and avoid the drift of the model. Experimental results show that the area under the tracking performance curve (AUC) score of the improved algorithm based on four benchmarks achieves at least 0.8% higher than C-COT, and 13% higher at most. This shows that the RFCG can effectively improve the tracking performance of the system.

Key words: C-COT; factorized convolution; over-fitting; compact generative model

1 引言

在机器视觉领域里,目标跟踪是一个非常基本的问题,并且在众多应用领域(如监控安防系统、无人机系统、无人驾驶系统、智能交通管理系统、人机交互系统,在线视觉跟踪系统)中,扮演着十分重要的角色。目标跟踪的基本原理是:给定目标物初始位置,结合一系列连续的图像帧,估算出目标物的运动轨迹。在线视觉跟踪的“在线”特点决定了,即使机器视觉对实时计算的约束条件异常复杂,理想的跟踪算法都应该保证系统的“精确性”和“鲁棒性”。

近些年来,可辨识相关分类器(DCF)在实时跟踪领域中扮演的角色越来越重要。国内外为了改善和提高基于 DCF 的改进算法的跟踪性能,基本上都是围绕向多维空间拓展^[1-2]、鲁棒性估计模型^[3-4]、机器学习模型^[5-6]、引入非线性基核函数法^[7]、应对边界效应策略^[8-9]几个角度展开研究的。但是,这些改进算法在提升系统跟踪性能的同时,图像的特征维度也随之提高,这一变化增加了系统的计算复杂程度和延迟。例如,Bolme 团队开发的第一代 MOSSE 算法^[10]比连续域卷积操作跟踪(C-COT)算法^[11]的计算速度快了近 1 000 倍^[12],但在精确性方面,后者只比前者提高一倍,同时 C-COT 算法引入了更多的约束条件,极易引起过拟合。

基于 C-COT 算法存在的问题,本文从模型维度、训练集大小和模板更新策略 3 个方面入手,在不牺牲跟踪系统性能的同时,降低跟踪系统的延迟和过拟合出现的概率。

2 C-COT 算法原理分析

2.1 构造连续域插值函数

定义一个图像的训练样本 $x_j \in \mathcal{X}$, $x_j^1, \dots$,

x_j^D 是它的 D 个特征通道, N_d 是 $\{x_j^d\}_{d=1}^D$ 的特征点数。构造差值函数如下:

$$J_d\{x^d\}(t) = \sum_{n=0}^{N_d-1} x^d[n] b_d\left(t - \frac{T}{N}n\right), \quad (1)$$

式(1)中, b_d 是差值函数, $t \in [0, T)$ 是连续域变量, T 是连续域支撑空间长度。插值函数 $J_d\{x^d\}(t)$ 由内插函数 b_d 的各个平移形式叠加而成,其中特征值 $x^d[n]$ 充当加权值。

2.2 构造置信函数

本文在连续区间 $[0, T)$ 上,定义置信函数 $s(t) = S_f\{x\}(t)$, 其中 $S_f: \mathcal{X} \rightarrow L^2(t)$ 是线性卷积算子。在连续域方程中, S_f 是由一系列分类器 $f = (f^1, \dots, f^D)$ 参数化得到的,其中 $\{f^d\}_1^D$ 是对应特征通道 d 的连续分类器。定义卷积算子,如下:

$$S_f\{x^d\} = \sum_{d=0}^D f^d * J_d\{x_d\}. \quad (2)$$

每个特征通道先用式(1)进行差值操作,然后和对应的分类器进行卷积运算,最后把分类器的卷积响应累加得到式(2)中的置信函数。

2.3 期望泛函

定义 y_j 是卷积算子 $S_f\{x_j\}$ 应用于训练样本 x_j 的期望输出,通过 m 个给定样本 $\{x_j, y_j\}_1^m$ 训练分类器,构造损失函数如下:

$$E(f) = \sum_{j=1}^m \alpha_j \|S_f\{x_j\} - y_j\|^2 + \sum_{d=1}^D \|\omega f\|^2, \quad (3)$$

式(3)中,权重 $\alpha_j \geq 0$ 控制着每个训练样本的影响力,惩罚系数 $\omega \geq 0$ 由特征点在整幅图像中的位置决定。样本处在背景区则 ω 值较大,在目标区则 ω 值较小。

2.4 训练连续分类器

本文通过最小化损失函数(3)进行分类器训练。式(1)的傅里叶系数 $\hat{J}_d\{x^d\}[k] = X^d[k] \hat{b}_d[k]$, 其中 $X^d[k] = \sum_{n=0}^{N_d-1} x^d[n] e^{-i\frac{2\pi}{N_d}nk}$, $k \in \mathbb{Z}$ 为 x^d 的离散傅里叶变换。同理,置信函数的离散傅

里叶系数如下:

$$\hat{S}_f\{x\}[k] = \sum_{d=1}^D \hat{f}^d X^d \hat{b}_d[k], \quad (4)$$

根据帕赛瓦尔(Parseval)公式,结合式(3)和(4)可得:

$$E(f) = \sum_{j=1}^m \alpha_j \left\| \sum_{d=1}^D \hat{f}^d X^d \hat{b}_d - \hat{y}_j \right\|_{\ell^2}^2 + \sum_{d=1}^D \left\| \omega f \right\|_{\ell^2}^2, \quad (5)$$

因此损失函数的最小化问题转化为分类器 f^d 的傅里叶系数 $\hat{f}^d[k]$ 的最小化问题。

2.5 正规方程的构造

考虑实际情况,训练分类器需要一个有限参数集,设该集合 $V = \text{span}\{e_k\}_{k=1}^{K_1} \times \cdots \times \text{span}\{e_k\}_{k=K_D}^{K_D}$ 。同时规定 $\hat{f}^d[k]=0, |k| > K_d, K_d = \frac{N_d}{2}$, 因此问题转化为用傅里叶系数 $\{\hat{f}^d[k]\}_{-K_d}^{K_d}$ 求式(5)的最小值问题。

定义向量 $\hat{f}^d = (\hat{f}^d[-K_d], \dots, \hat{f}^d[K_d])^T$, 向量系 $\hat{f} = [(\hat{f}^1)^T, \dots, (\hat{f}^D)^T]^T$, 把 y_j 的 $(2K+1)$ 个傅里叶系数定义成向量 $\hat{y}_j = (y_j[-K], \dots, y_j[K])^T$, 其中 $K = \max\{x_d\}_{d=1}^D$ 。

定义 L 是 $\hat{\omega}[k]$ 中非零项的数量, 因此有 $\hat{\omega}[k]=0, |k| > 0$ 。定义托普利茨(Toeplitz)矩阵 W_d , 大小为 $(2K_d+2L+1) \times (2K_d+1)$, 对应的卷积算子为 $W_d \hat{f}_d = \text{vec} \hat{\omega} * \hat{f}^d$, 最后记块对角矩阵 $W = W_1 \oplus \cdots \oplus W_D$, 因此式(5)的最小化问题就转化为式(6)的平方和最小化问题, 如下:

$$E_V(\hat{f}) = \sum_{j=1}^m \alpha_j \left\| \sum_{d=1}^D A_d \hat{f} - \hat{y}_j \right\|_2^2 + \sum_{d=1}^D \left\| W_d \hat{f} \right\|_2^2, \quad (6)$$

式(6)中矩阵 $A_j = [A_j^1, \dots, A_j^D]$ 有 $(2K+1)$ 行, 对应特征通道 d 的块对角矩阵 A_j^d 的列数为 $(2K_d+1)$, 其中 $\{X_j^d[k] b_d[k]\}_{-K_d}^{K_d}$ 是矩阵 A_j^d 的构成元素。

为获取正规方程的简化表达式, 定义样本矩阵 $A = [A_1^T, \dots, A_m^T]$, 与之对应的权重对角矩阵 $\Gamma = \alpha_1 I \oplus \cdots \oplus \alpha_m I$, 标签向量 $\hat{y} = (y_1^T, \dots, y_m^T)^T$, 因此式(6)的最小值问题转化为求解正规方程(7)的问题, 如下所示:

$$(A^H \Gamma A + W^H W) \hat{f} = A^H \Gamma \hat{y}, \quad (7)$$

式(7)中 $(A^H \Gamma A + W^H W)$ 为对称正定方程, 可以用共轭梯度迭代(Conjugate Gradient)法求解。

3 C-COT 算法的改善

3.1 对 C-COT 算法进行改善的原因和目的

本文的目的是提高 DCF 算法的计算速度和跟踪性能。从 MOSSE 算法依次到 KCF 算法、DSST 算法、CN 算法、SRDCF 算法、C-COT 算法, 跟踪性能的确越来越好, 但模型越来越复杂, 计算速度也越来越慢。导致计算速度降低的原因主要有: 训练的模型维度越来越高、训练集越来越大、模板更新的速度越来越快。本文的目的是在不牺牲跟踪性能的同时, 降低跟踪系统的延迟和 C-COT 算法中过拟合出现的概率。

3.2 卷积因式分解

本文在 C-COT 算法的基础上, 对特征提取进行了简化, 使用 C-COT 算法的特征子集, 即在原来的 D 维特征中, 选取其中的 C ($C < D$) 维特征。C-COT 算法中的每个维度特征都对应一个分类器, 但多数分类器对系统的跟踪性能贡献很小, 所以本文只选择其中贡献较多的 C 个分类器 $f^1, \dots, f^C$, 因此每一维特征通道的分类器都可以用这 C 个分类器的线性组合 $\sum_{c=1}^C p_{d,c} f^c$ 表征。

定义新的卷积算子, 如下:

$$S_{pf}\{x\} = P f * J\{x\}, \quad (8)$$

其中 $P = (p_{d,c})$ 是 $D \times C$ 矩阵, 定义分类器向量 $f = (f^1, \dots, f^C)^T$, 标签函数 $\hat{y} = (y_1, \dots, y_C)^T$, 因此, 式(8)进行因式分解操作, 如下:

$$S_{pf}\{x\} = P f * J\{x\} =$$

$$\sum_{d,c} p_{d,c} f^c * J_d\{x^d\} = f * P^T J\{x\}, \quad (9)$$

因为式(9)满足卷积的线性性质, 故其因式分解可归纳为两步, 即包含 D 维特征的特征向量 $J\{x\}(t)$ 先在 t 处, 与降维矩阵 P^T 相乘, 得到包含 C 维特征的特征向量再与分类器向量 f 进行卷积运算。

本文通过最小化损失函数(10)获得分类器向量 f 和降维矩阵 P , 设定特征图 $z = J\{x\}$ 的特征

通道的离散傅里叶系数 $\hat{z}^d = X^d[k]\hat{b}_d[k]$, 类比值(5)的推导思想训练因式分解卷积, 得到式(9)对应的新损失函数, 如下:

$$E(f, \mathbf{P}) = \left\| \hat{\mathbf{z}}^T \mathbf{P} \hat{\mathbf{f}} - \hat{\mathbf{y}} \right\|_{\ell^2}^2 + \sum_{d=1}^C \left\| \hat{\omega} * \hat{f}^d \right\|_{\ell^2}^2 + \lambda \left\| \mathbf{P} \right\|_F^2, \quad (10)$$

式(10)中, 新增的 F 范数项 $\lambda \left\| \mathbf{P} \right\|_F^2$ 的作用是正则化降维矩阵 $\mathbf{P}$ 。不同于式(6), 式(10)要解决的是非线性最小二乘问题。由 $\hat{\mathbf{z}}^T \mathbf{P} \hat{\mathbf{f}}$ 的双线性特点决定, 该问题又转化为矩阵因式分解问题^[21]。由参数的大小问题和系统的在线特性决定, 如“交替最小二乘法”一类的方法在处理最

优解时不可取, 所以本文采用高斯-牛顿 (Gauss-Newton) 法^[32] 和共轭梯度法求解“二次子问题”。Gauss-Newton 法的思想是: 利用一阶泰勒级数展开式最小化式(10)的残差平方和。本文以 $(\hat{f}_i, \mathbf{P}_i)$ 为基准, 对双线性项 $\hat{\mathbf{z}}^T \mathbf{P} \hat{\mathbf{f}}$ 进行近似估计, 如下:

$$\hat{\mathbf{z}}^T (\mathbf{P} + \Delta \mathbf{P}) \hat{\mathbf{f}}_{i,\Delta} \approx \hat{\mathbf{z}}^T \mathbf{P}_i \hat{\mathbf{f}}_{i,\Delta} + (\hat{\mathbf{f}} \otimes \hat{\mathbf{z}})^T \text{vec}(\Delta \mathbf{P}), \quad (11)$$

式(11)中 $\hat{\mathbf{f}}_{i,\Delta} = (\hat{f}_i + \Delta \hat{f})$, 算子“ $\otimes$ ”是克罗内科 (Kronecker) 积算子, 用于获得增量矩阵 $\Delta \mathbf{P}$ 。把式(11)带入式(10)中, 可得高斯-牛顿法第 i 次迭代结果如下:

$$\tilde{E}(\hat{f}_{i,\Delta}, \Delta \mathbf{P}) = \left\| \hat{\mathbf{z}}^T \mathbf{P} \hat{\mathbf{f}} + (\hat{\mathbf{f}} \otimes \hat{\mathbf{z}})^T \text{vec}(\Delta \mathbf{P}) - \hat{\mathbf{y}} \right\|_2^2 + \sum_{d=1}^C \left\| \hat{\omega} * \hat{f}^d \right\|_2^2 + \mu \left\| \mathbf{P} + \Delta \mathbf{P} \right\|_F^2. \quad (12)$$

因本文设定分类器 f 的非零傅里叶系数是有限的, 所以式(12)是最小二乘问题, 可参照正规方程(7)进行求解, 不同的是附加了一项增量矩阵 $\Delta \mathbf{P}$ 。本文利用 CG 法对 Gauss-Newton 法中二次子问题进行优化, 得到了全新的分类器 $\hat{f}_{i,\Delta}^*$ 和增量矩阵 $\Delta \mathbf{P}^*$, 可对式(9)中的分类器进行更新: $\hat{f}_{i+1} = \hat{f}_{i,\Delta}^*$, $\mathbf{P}_{i+1} = \mathbf{P}_i + \Delta \mathbf{P}^*$ 。

卷积因式分解的主要目的是, 降低特征通道的维数, 进而降低跟踪系统的计算复杂程度和内存空间占用率。由分类器的自适应特点决定, 降维矩阵 $\mathbf{P}$ 可以通过训练第一帧图像获得, 并在后续跟踪过程中保持不变, 故本文仅通过存储预测值 $\mathbf{P}^T J \{x_j\}$ 的方式, 就节省了大量内存空间, 且 $\mathbf{P}^T J \{x_j\}$ 还可作为 C-COT 算法的输入, 经训练能得到一个 C 维分类器, 即将特征维数 D 降低到 C , 进而降低系统计算复杂程度。

3.3 紧凑衍生模型

在 C-COT 算法中, 跟踪系统刻画一个目标物的简单运动就需要大量样本, 而且其中很多样本是冗余样本, 这点极易造成过拟合。

本文提出一个更具代表性的小模型, 即紧凑衍生模型, 既可以避免存储完整的样本训练集, 又能提高样本的多样性。本文选择的训练模型取自服从高斯分布的样本集, 这样既能保证每个样本代表特定的运动时期, 又能大大降低过拟合出现的风险。

紧凑衍生模型的原理是: 为排除冗余样本, 引

入联合概率密度函数 $p(x, y) = \sum_{j=1}^{M_{\max}} \pi_j \delta_{x_j, y_j}(x, y)$, 其中 x 是样本特征图, y 是相对应的标签函数, 权重 $\pi_j \sim (1 - \gamma)^{M_{\max} - j}$, $\gamma \in [0, 1]$ 是学习速率, δ_{x_j, y_j} 是狄拉克 (Dirac) 脉冲函数, $M_{\max}$ 为样本上限。因为标签函数 y 只由样本 x 距目标中心的远近程度决定, 故设定目标中心处 $y = y_0$, 非目标中心处 $y = 0$, 联合概率密度函数可进一步更新为 $p(x, y) = p(x) \delta_{y_0}(y)$, 本文只需对 $p(x)$ 进行估计即可。用期望项替代式(5)中的累加项, 可得损失函数如下:

$$E(f) = E \left\{ \left\| S_x^f - y \right\|_2^2 \right\} + \sum_{d=1}^D \left\| \omega f^d \right\|_{\ell^2}^2, \quad (13)$$

式(13)中期望项 $E \left\{ \left\| S_x^f - y \right\|_{\ell^2}^2 \right\}$ 仅由 $p(x, y)$ 决定, 且样本集 $\{x_j\}_1^{M_{\max}}$ 服从高斯混合分布, 有:

$$P(x) = \sum_{l=1}^L \pi_l N(x, \mu_l), \quad (14)$$

式(14)中样本的权重 π_l 满足 $\sum_{l=1}^L \pi_l = 1$, $0 \leq \pi_l \leq 1$, μ_l 是样本的均值。为更新高斯混合分布, 设定初始分量: $\pi_0 = \gamma$, $\mu_0 = x_j$, $L = M_{\max}$ (后续实验中, 设定 $M_{\max} = \frac{m}{8}$)。规定每当有新样本到来, 样本集都重新归类。若样本容量未满, 新来的样

本直接被收纳;若样本集已满,如有权重 π_j 低于阈值的样本,直接剔除该样本并收纳新样本,如没有则将新样本加入样本集,而后比较这 $(M_{\max} + 1)$ 个样本中权重大小,并将权重最接近的两个样本合并,如下:

$$\begin{cases} \pi_n = \pi_k + \pi_l \\ \mu_n = \frac{\pi_k \mu_k + \pi_l \mu_l}{\pi_k + \pi_l} \end{cases}, \quad (15)$$

结合式(15),式(13)可转化为式(16):

$$E(f) = \sum_{l=1}^{M_{\max}} \pi_l \left\| S_{\mu_l}^f - y_0 \right\|_{\ell^2}^2 + \sum_{d=1}^D \left\| \omega f^d \right\|_{\ell^2}^2, \quad (16)$$

式(16)可参考式(7)求解分类器。

3.4 模型更新策略

如果为求取求解分类器,对图像的每一帧都进行迭代,不仅会影响跟踪系统的计算速度,还容易造成过拟合。因此本文每隔 $N_s = 5$ 帧进行一次模型的更新,即执行式(7)的迭代求解分类器。因为式(7)是动态变化的,不必精确求解,所以仅在处理第一帧图像时,设定 CG 法迭代次数 $N_{CG} = 100$,而在处理后续帧时,无需过多迭代,设定 $N_{CG} = 5$ 。基于这种迭代方式,平均每帧只进行 N_{CG}/N_s 次迭代,极大地提升了运行效率。同时,样本集的更新,则是每一帧都进行。

4 实验与结果

本文基于 4 个标准数据集,通过综合实验验证了本文提出的改善方法。这 4 个标准数据集分别是:VOT2016 数据集^[12]、UAV123 数据集^[13]、OTB-2015 数据集^[14]以及 Temple-Color 数据集^[15]。

4.1 实验准备

本文基于 Matlab 环境进行算法验证,使用与 C-COT 算法相同的 4 种特征提取方式,即在 VGG-m 网络^[16]中的第一个卷积层(conv1)和最后一个卷积层(conv5)进行特征提取,及 HOG 法^[17]和 Color Names (CN)法^[18]这两种手动特征提取方式。通过这 4 种不同的特征提取方式训练因式分解卷积方法中的降维矩阵和分类器,结果如表 1 所示。

设定式(10)中的正则化参数 $\lambda = 2 \times 10^{-7}$,第

表 1 在不同特征提取条件下的特征维度 D 和分类器数目 C

Tab.1 Feature dimensionality D and the number of filters C for each features

	Conv-1	Conv-5	HOG	CN
特征维度 D	96	512	31	11
分类器纬度 C	16	64	10	3

一次迭代的初值 $i = 0$, $\hat{f}_i = 0$,结合式(12)对式(10)分别进行了 10 次 Gauss-Newton 迭代和 20 次 CG 迭代。为了保证跟踪的精确性,本文使用 PCA 初始化系数矩阵 P_0 (与随机参数初始化鲁棒性相同)。

本文设定紧凑衍生模型学习速率 $\gamma = 0.012$,样本的数量 $L = 50$,与 C-COT 算法使用的样本数($m = 400$)相比减少了 80%。模型更新时,每 $N_s = 6$ 帧更新一次分类器(与 C-COT 算法一致),CG 迭代次数 $N_{CG} = 5$ 。同时,所有视频数据集的全部参数设置固定不变。

4.2 与 C-COT 算法的对比与分析

本章节基于 VOT2016 数据集,与 C-COT 算法进行对比分析,VOT2016 数据集包含 60 个典型视频。使用预期平均重叠率(EAO)对总体性能进行评估,该评估兼顾精确性和鲁棒性^[19]。

表 2 本文提出的改善算法做出的贡献

Tab.2 Contributions of improvement strategy

跟踪 算法	C-COT	因式分 解卷积	紧凑衍 生模型	模板更 新策略
EAO	0.331	0.342	0.352	0.374
FPS	0.3	1.1	2.6	6.0
维度变化	-	$D \rightarrow C$	$M \rightarrow L$	$N_{CG} \rightarrow \frac{N_{CG}}{N_s}$
计算量减少率	—	$6 \times$	$8 \times$	$6 \times$

表 2 显示了本文在 C-COT 算法改善方面做出的贡献。把因式分解卷积集成到 C-COT 算法中,使得 EAO 指数提高至 0.342,系统计算量减少了 6 倍。在此基础上,紧凑衍生模型将 EAO 指数提高至 0.352,计算量减少了 8 倍。模板更新策略进一步将 EAO 指数提高到 0.374,计算量减少了 6 倍。

本文在单核 CPU 上进行实验, FPS 表征系统每秒可以处理的图像帧数。表 2 还展示了改善算法对跟踪系统运行速度的提升(不考虑特征提取占用的时间);改善算法的 3 项工作都不同程度地提高了系统的跟踪速度,相比与 C-COT 算法,速度一共提升了近 20 倍。

本文在不改变原算法跟踪精度的前提下,显著地提升了跟踪速度,这对改善跟踪系统的实时性有很大帮助。

4.3 与主流跟踪算法的对比与分析

本文基于 4 个标准数据集,与最新的主流跟踪系统进行比较。

4.3.1 基于 VOT2016 数据集的对比与分析

本文从 EAO 指数、鲁棒性、精确性和运行速度(这里考虑特征提取占用的时间)4 个方面,分别与最新的主流跟踪系统进行比较。在 EAO 指数方面, VOT2016 挑战赛中排名第一的 C-COT 算法的 EAO 指数为 0.331, 本文与之相比, EAO 指数增加了 13.0%;在精确性方面, 本文达到了 0.72 的最低失误差率;在运行速度方面, 本文仅采用手动特征提取(HOG 特征和 CN 特征)的方式就达到了最快运行速度, 详见表 3。

表 3 不同跟踪策略的实验结果对比与分析

Tab.3 Compare and analysis of experimental results from different tracking strategies

跟踪算法	识别率	误识别率	精确度	处理速度
SRBT	0.290	1.25	0.50	3.69
EBT	0.291	0.90	0.44	3.01
DDC	0.293	1.23	0.53	0.20
Staple	0.295	1.35	0.54	11.14
MLDF	0.311	0.83	0.48	1.48
SSAT	0.321	1.04	0.57	0.48
TCCN	0.325	0.96	0.54	1.05
C-COT	0.331	0.85	0.52	0.51
RDFC&CGM*	0.322	1.08	0.53	15.13
RDFC*	0.374	0.72	0.54	4.53

本文对跟踪速度的提升不是通过缩短特征提取时间实现的,而是通过有效降低图像处理过程中的计算时间实现的。

4.3.2 基于 UAV123 数据集的对比与分析

“无人机空中跟踪技术”近年来受到广泛关注,并且在野外生命监测、搜救、导航、监控等领域都有广泛的应用,因此无人机持续导航以及实时跟踪技术成为关键。基于此应用背景,理想的跟踪系统应该是精确且稳定的,并且能够在有限的硬件环境下实时运行。此环节采用 HOG 特征和 CN 特征进行手动特征提取,且在 i7 单核 CPU 上以每秒 60 帧的速度运行。

UAV123 数据集由 123 个最新航拍视频组成,帧数超过 11 万,采用“跟踪性能曲线”第一象限积分(AUC)大小来评定跟踪算法的性能。

图 1(a)是以 UAV123 数据集为基础,模拟出的跟踪性能曲线。在排名前五的跟踪系统中,只有基于 Staple 算法的跟踪系统是实时运行的, AUC 得分为 45.3%。在帧频不超过 60 的情况下,基于本文改善算法(RFCG & CN)的跟踪系统也是实时运行的,且 AUC 得分为 51.7%,比 Staple 算法得分高出 6.4%。与 C-COT 算法相比,本文改善算法(RFCG)的跟踪性能略占优势, AUC 得分为 53.7%。

本文虽然显著提高了 C-COT 跟踪算法的跟踪速度,但是在众多算法中,跟踪速度不是最快的。但是在跟踪速度较快的一系列算法中,跟踪精度是比较高的。

4.3.3 基于 OTB2015 数据集的对比与分析

本文与 20 个先进的跟踪算法进行对比,它们分别是: TLD^[20]、Struck^[21]、CFLB^[9]、ACT^[1]、TGPR^[22]、KCF^[7]、DSST^[3]、SAMP^[23]、MEEM^[24]、DAT^[25]、LCT^[26]、HCF^[27]、SRDCF^[28]、SRDCFad^[29]、DeepSRDCF^[30]、Staple^[31]、MDNet^[32]、SiameseFC^[33]、TCNN^[34]、C-COT^[11]。图 1(b)是以 OTB2015 数据集中的 100 个视频为基础,模拟出的跟踪性能曲线^[35-37]。在手动特征提取的条件下,进行跟踪系统性能比较。SRDCFad 算法的 AUC 得分较高,为 63.4%,而本文改善算法(RFCG & CN)的跟踪性能略占优势, AUC 得分为 65.0%。在使用单核 CPU,视频每秒 60 帧的前提条件下, C-COT、MDNet 和 TCNN 这 3 类算法的 AUC 得分较高,分别为 69.0%、68.5% 和 66.1%,而本文改善算法(RFCG)的跟踪性能最优, AUC 得分为 70.0%。

本文在提高原有跟踪算法速度的基础上,也

在一定程度上挺高了跟踪精度。

4.3.4 基于 Temple-Color 数据集的比较与分析

图 1(c)是以 Temple-Color 数据集中 128 个视频为基础,模拟出的跟踪性能曲线。本文的改

善算法(RDFC)较之 C-COT 算法,跟踪性能有了一定的提升,AUC 得分提高了 0.8%。

本文基于针对色彩空间的样本,在跟踪精度方面的有一定改善,但不明显。

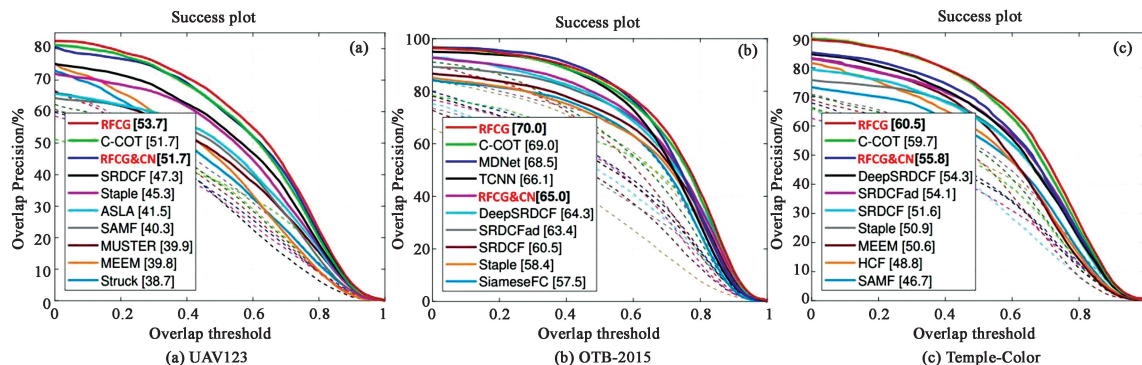


图 1 三类数据集下的跟踪性能曲线

Fig. 1 Trace performance curves under three types of datasets

5 结 论

本文以 DCF 方法为基础,在 C-COT 算法的基础上,引入了一个因式分解卷积算子,降低了模型中维度,提出了一个训练样本的紧凑衍生模型,在提高样本多样性的同时,降低了系统计算时间和过拟合出现的概率。最后,提出了一个简单有效的模型更新策略,进一步降低了过拟合出现的概率。实验结果证明:基于 VOT2016 数据集,在 EAO 指数方面,本文获得的分数为 37.4%,比 C-COT 算法高出 13.0%,

在精确性方面,本文达到了 0.72 的最低失误率;基于 UAV123 数据集,本文获得了 53.7% 的 AUC 分数,比 C-COT 算法高出 2%;基于 OTB2015 数据集,本文获得了 70.0% 的 AUC 分数,比 C-COT 算法高出 1%;基于 Temple-Color 数据集,本文获得了 60.5% 的 AUC 分数,比 C-COT 算法高出 1%。

从跟踪性能角度分析,本文在一定程度上改善了系统跟踪精度,明显地提高了原跟踪算法的跟踪速度。跟踪速度的提高降低了系统的硬件标准,这使得复杂的跟踪算法可以应用到更多的实际项目中去。

参 考 文 献:

- [1] DANELLJAN M, KHAN F S, FELSBERG M, *et al.* Adaptive color attributes for real-time visual tracking [C]// *Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition*. Columbus, OH, USA: IEEE, 2014.
- [2] GALOOGAHI H K, SIM T, LUCEY S. Multi-channel correlation filters [C]// *Proceedings of 2013 IEEE International Conference on Computer Vision*. Sydney, NSW, Australia: IEEE, 2014.
- [3] WANG R L, ZOU J C, CHE M Q, *et al.* Robust and real-time visual tracking based on single-layer convolutional features and accurate scale estimation [C]// *Proceedings of the 13th Conference on Image and Graphics Technologies and Applications*. Beijing, China: Springer, 2018: 471-482.
- [4] DANELLJAN M, HÄGER G, KHAN F S, *et al.* Discriminative scale space tracking [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(8): 1561-1575.
- [5] BIBI A, MUELLER M, GHANEM B. Target Response Adaptation for Correlation Filter Tracking [M]. LEIBE B, MATAS J, SEBE N, *et al.* *Computer Vision-ECCV 2016*. Cham: Springer, 2016.
- [6] DANELLJAN M, HÄGER G, KHAN F S, *et al.* Adaptive decontamination of the training set: a unified formula-

- tion for discriminative visual tracking [C]//*Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, NV, USA; IEEE, 2016.
- [7] 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.
- [8] DANELLJAN M, HÄGER G, KHAN F S, *et al.* Learning spatially regularized correlation filters for visual tracking [C]//*Proceedings of 2015 IEEE International Conference on Computer Vision*. Santiago, Chile; IEEE, 2016.
- [9] GALOOGAHI H K, SIM T, LUCEY S. Correlation filters with limited boundaries [C]//*Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition*. Boston, MA, USA; IEEE, 2015.
- [10] BOLME D S, BEVERIDGE J R, DRAPER B A, *et al.* Visual object tracking using adaptive correlation filters [C]//*Proceedings of 2010 IEEE Computer Society IEEE Conference on Computer Vision and Pattern Recognition*. San Francisco, CA, USA; IEEE, 2010.
- [11] DANELLJAN M, ROBINSON A, KHAN F S, *et al.* Beyond correlation filters: learning continuous convolution operators for visual tracking [C]//*Proceedings of the 14th European Conference on Computer Vision*. Amsterdam, The Netherlands; Springer, 2016.
- [12] KRISTAN M, LEONARDIS A, MATAS J, *et al.* The Visual Object Tracking VOT2016 Challenge Results [M]. HUA G, JÉGOU H. *Computer Vision-ECCV 2016 Workshops*. Cham; Springer, 2016.
- [13] MUELLER M, SMITH N, GHANEM B. A benchmark and simulator for UAV tracking [C]//*Proceedings of the 14th European Conference on Computer Vision*. Amsterdam, The Netherlands; Springer, 2016; 445-461.
- [14] WU Y, LIM J, YANG M H. Object tracking benchmark [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(9): 1834-1848.
- [15] LIANG P P, BLASCH E, LING H B. Encoding color information for visual tracking: algorithms and benchmark [J]. *IEEE Transactions on Image Processing*, 2015, 24(12): 5630-5644.
- [16] DAM N, NGUYEN V T, DO M N, *et al.* Realtime Face Verification with Lightweight Convolutional Neural Networks [M]. BEBIS G, BOYLE R, PARVIN B, *et al.* *Advances in Visual Computing*. Cham; Springer, 2015.
- [17] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection [C]//*Proceedings of 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. San Diego, CA, USA; IEEE, 2005.
- [18] VAN DE WEIJER J, SCHMID C, VERBEEK J, *et al.* Learning color names for real-world applications [J]. *IEEE Transactions on Image Processing*, 2009, 18(7): 1512-1523.
- [19] FELSBERG M, BERG A, HAGER G, *et al.* The thermal infrared visual object tracking VOT-TIR2015 challenge results [C]//*Proceedings of 2015 IEEE International Conference on Computer Vision Workshop*. Santiago, Chile; IEEE, 2015.
- [20] KALAL Z, MATAS J, MIKOLAJCZYK K. P-N learning: bootstrapping binary classifiers by structural constraints [C]//*Proceedings of 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. San Francisco, CA, USA; IEEE, 2010.
- [21] HARE S, SAFFARI A, TORR P H S. Struck: structured output tracking with kernels [C]//*Proceedings of 2011 International Conference on Computer Vision*. Barcelona, Spain; IEEE, 2011.
- [22] GAO J, LING H B, HU W M, *et al.* Transfer learning based visual tracking with Gaussian processes regression [C]//*Proceedings of the 13th European Conference on Computer Vision*. Zurich, Switzerland; Springer, 2014.
- [23] LI Y, ZHU J K. A scale adaptive kernel correlation filter tracker with feature integration [C]//*Proceedings of the European Conference on Computer Vision*. Zurich, Switzerland; Springer, 2014.
- [24] ZHANG J M, MA S G, SCLAROFF S. MEEM: robust tracking via multiple experts using entropy minimization [C]//*Proceedings of the 13th European Conference on Computer Vision*. Zurich, Switzerland; Springer, 2014.
- [25] POSSEGER H, MAUTHNER T, BISCHOF H. In defense of color-based model-free tracking [C]//*Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition*. Boston, MA, USA; IEEE, 2015.
- [26] MA C, YANG X K, ZHANG C Y, *et al.* Long-term correlation tracking [C]//*Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition*. Boston, MA, USA; IEEE, 2015.

- [27] MA C, HUANG J B, YANG X K, *et al.* Hierarchical convolutional features for visual tracking [C]//*Proceedings of 2015 IEEE International Conference on Computer Vision*. Santiago, Chile; IEEE, 2015.
- [28] DECLERCQ A, PIATER J H. On-line simultaneous learning and tracking of visual feature graphs [C]//*Proceedings of 2007 IEEE Conference on Computer Vision and Pattern Recognition*. Minneapolis, MN, USA; IEEE, 2007.
- [29] HONG J H, FITZGIBBON A. Secrets of matrix factorization; approximations, numerics, manifold optimization and random restarts [C]//*Proceedings of 2015 IEEE International Conference on Computer Vision*. Santiago, Chile; IEEE, 2015.
- [30] DANELLJAN M, HÄGER G, KHAN F S, *et al.* Convolutional features for correlation filter based visual tracking [C]//*Proceedings of 2015 IEEE International Conference on Computer Vision Workshop*. Santiago, Chile; IEEE, 2015.
- [31] BERTINETTO L, VALMADRE J, GOLODETZ S, *et al.* Staple: complementary learners for real-time tracking [C]//*Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, NV, USA; IEEE, 2016.
- [32] NAM H, HAN B. Learning multi-domain convolutional neural networks for visual tracking [C]//*Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, NV, USA; IEEE, 2016.
- [33] BERTINETTO L, VALMADRE J, HENRIQUES J F, *et al.* Fully-convolutional Siamese networks for object tracking [C]//*Proceedings of the European Conference on Computer Vision*. Amsterdam, The Netherlands; Springer, 2016.
- [34] NAM H, BAEK M, HAN B. Modeling and propagating CNNs in a tree structure for visual tracking [J]. *arXiv preprint arXiv:1608.07242*, 2016.
- [35] 闫辉,许廷发,吴青青,等.多特征融合匹配的多目标跟踪[J].中国光学,2013,6(2):163-170.
YAN H, XU T F, WU Q Q, *et al.* Multi-object tracking based on multi-feature joint matching [J]. *Chinese Optics*, 2013, 6(2): 163-170. (in Chinese)
- [36] 王铭明,陈涛,王建立,等.Mean-shift 跟踪算法及其在光电跟踪系统中的应用[J].中国光学,2014,7(2):332-338.
WANG M M, CHEN T, WANG J L, *et al.* Mean-shift tracking algorithm and its application in optoelectronic tracking system [J]. *Chinese Optics*, 2014, 7(2): 332-338. (in Chinese)
- [37] 江山,张锐,韩广良,等.复杂背景灰度图像下的多特征融合运动目标跟踪[J].中国光学,2016,9(3):320-328.
JIANG S, ZHANG R, HAN G L, *et al.* Moving object tracking based on multi-feature fusion in the complex background gray image [J]. *Chinese Optics*, 2016, 9(3): 320-328. (in Chinese)

作者简介:



王 鑫(1993—),男,黑龙江绥化人,硕士,2016 年于哈尔滨工业大学获得硕士学位,主要从事图像跟踪方面的研究。E-mail: 14S104020@hit.edu.cn