

doi:10.11830/ISSN.1000-5013.201701024



结合 PSO 的改进压缩跟踪方法

刘 韶 涛, 姚 灿 荣

(华侨大学 计算机科学与技术学院, 福建 厦门 361021)

摘要: 针对基于在线检测的跟踪方法中目标在多尺度空间中的搜索和匹配问题,结合粒子群优化算法(PSO)和压缩感知思想,提出一种鲁棒的多尺度目标跟踪算法.首先,通过粒子群在多尺度空间中采集样本;然后,经过压缩感知提取特征;最后,通过粒子的迭代计算,搜索出当前目标的最佳匹配位置.实验结果表明:提出的算法能较好地适应目标的多尺度变化,在快速性和鲁棒性上具有更好的性能.

关键词: 目标跟踪; 压缩感知; 粒子群优化; 多尺度

中图分类号: TP 391.3 **文献标志码:** A **文章编号:** 1000-5013(2017)01-0121-06

Improved Compress Tracking Algorithm Based on PSO

LIU Shaotao, YAO Canrong

(College of Computer Science and Technology, Huaqiao University, Xiamen 361021, China)

Abstract: For the searching and matching problem in multi-scale space of online detecting tracking method, a robust multi-scale tracking algorithm was proposed based on particle swarm optimization (PSO) and compress sensing. Firstly, feature was sampled with particles in multi-scale space. Then feature was extracted by compress sensing. Finally, targets would be searched quickly and robustly after calculate the best fitness and position of all the particle. The experimental results demonstrate that the proposed algorithm can adapt target in multi-scale change and has a better performance in robustness and rapidity.

Keywords: visual tracking; compress sensing; particle swarm optimization; multi-scale

一个典型的目标跟踪系统包括表观模型和在当前帧定位目标新位置的搜索策略^[1].由于光照变化、遮挡等问题,导致目标的表观模型时常发生变化.因此,一个自适应强的表观模型表示对鲁棒的目标跟踪十分重要.基于特征实时检测^[2-3]的运动目标跟踪方法是近年来的一种主流目标跟踪技术^[4-6].在线方法在对目标进行分类判别的同时,也在对图像进行采样,用于实时更新分类器,使检测器具有一定的适应能力,不仅在一定程度上解决了检测和跟踪时的遮挡问题,还能适应复杂场景下的多目标跟踪.然而,其计算和搜索复杂度较高,且容易产生漂移.近年来,基于随机优化技术的粒子群优化算法(PSO)被广泛应用于计算机视觉处理^[7-9].为了在多尺度空间中快速寻找目标,降低计算代价,保持跟踪的实时性和重叠率,本文提出一种结合粒子群优化的压缩跟踪算法(PSO-CT),在每一帧的目标搜索过程中,利用粒子群优化的搜索方法,在多尺度空间中更快速地搜索目标.

1 压缩跟踪算法

压缩跟踪(compress tracking)算法是 Zhang 等^[6]基于压缩感知理论^[10]提出的一种高效、快速的基

于在线学习的实时跟踪方法. 压缩跟踪算法的流程, 如图 1 所示. 图 1 中: 分类器更新是通过在当前帧跟踪结果的邻域内采集正负样本, 然后进行压缩, 训练分类器得到.

1.1 随机投影

如果信号通过某种变换后是可稀疏表示或者可压缩的, 则可以设计一个与变换基不相关的测量矩阵来测量信号, 得到的测量值通过求解优化问题可实现信号的精确或者近似重构. 具体地, 对于高维图像空间中的数据 $x \in \mathbb{Z}$, 如果 x 在正交基 Ψ 上是 K -稀疏的, 且能找到一个与 Ψ 不相关的观测基 R , 则可以使用随机观测矩阵 $R \in \mathbb{Z}^{n \times m}$, 将其投影到低维空间 $y \in \mathbb{Z}^n$ 上, 如图 2 所示. 图 2 中: $n \ll m$. Johnson-Lindenstrauss 准则表明: 当随机测量矩阵 R 满足有限等距性质时, K 个系数可通过求解最小误差, 从 y 中高精度地恢复原始信号 x .

1.2 随机测量矩阵

为了适应实际应用的实时性需求, 采用文献[6]提出的一个非常稀疏的随机测量矩阵, 具体矩阵元素定义为

$$r_{i,j} = \sqrt{s} \times \begin{cases} 1, & p = 1/2s, \\ 0, & p = 1 - (1/s), \\ -1, & p = 1/2s. \end{cases} \quad (1)$$

式(1)中: p 为随机元素 $r_{i,j}$ 取某个值的概率. 例如, 取 $s = n/4$ 时, 测量矩阵非常稀疏, 即 n 行中每行只需 $2 \sim 4$ 个非零元素, 就能满足 Johnson-Lindenstrauss 推论, 且只需初始的一次离线计算就能满足整个跟踪过程的计算.

假设 s 是第 $t-1$ 跟踪到的目标位置, 在以 s 为中心, r 为半径的范围内, 采集测试样本集的 Haar-like 特征: $T_i \in \mathbb{Z}^{w \times h \times l} (i=1, 2, \dots, r^2)$, l 为 Haar-like 的模板数, 则每个特征的维数 m 可能达到 10^6 以上. 利用初始产生的随机稀疏矩阵, 将此高维数据投影到低维空间中, 形成适合快速计算的低维特征 v .

1.3 目标的 Haar-like 多尺度表示与尺度不变性

对于每个样本 $X \in \mathbb{Z}^{w \times h}$, 它的多尺度表示是多个矩形滤波器在多个尺度 $\{F_{1,1}, F_{1,2}, \dots, F_{1,w \times h}\}$ 组成, 如图 3 所示. 即

$$F_{w \times h}(x, y) = \frac{1}{w \times h} \times \begin{cases} 1, & 1 \leq x \leq w, \quad 1 \leq y \leq h, \\ 0, & \text{其他.} \end{cases} \quad (2)$$

式(2)中: w, h 分别为矩形滤波器的宽和高. 用一个 $w \times h$ 维的向量表示特征, 再用高维特征串联这些向量, 形成维数为 $m = (w \times h)^2$ 的向量, 这个向量维数通常能达到 10^6 以上. 之后, 将特征 x 投影到一个稀疏随机矩阵 Z , 得到一个低维特征 $v, v \in \mathbb{Z}^n$. 由文献[6]可知, 低维的 Haar-like 特征向量具有尺度不变性.

2 基于 PSO 改进的压缩跟踪算法

2.1 粒子群优化算法

一个粒子群优化算法代表某个种群^[11], 每个粒子代表问题的一个潜在解. 每次迭代时, 粒子在一个 n 维搜索空间移动, 并将与其相关联的位置矢量 $x_i^t = \{x_{i,1}^t, x_{i,2}^t, \dots, x_{i,n}^t\}$, 速度矢量 $v_i^t = \{v_{i,1}^t, v_{i,2}^t, \dots, v_{i,n}^t\}$ 进行迭代进化. 粒子更新示意图, 如图 4 所示.

在搜索空间里, PSO 的每个粒子的移动依赖于个体认知 p 和社会认知 g . 个体认知代表粒子个体历史移动过程中的最佳位置; 社会认知代表整个粒子群在搜索过程中的最佳位置. 通过整合这 2 种类型的

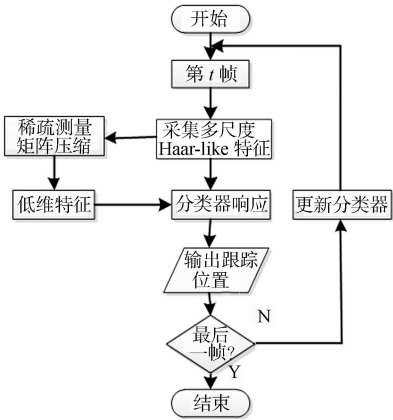


图 1 实时压缩跟踪算法流程
Fig. 1 Real-time compress tracking algorithm

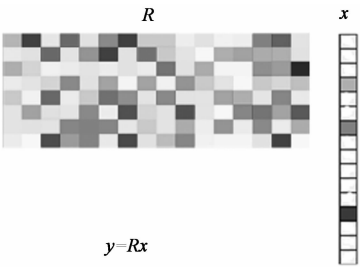


图 2 特征映射
Fig. 2 Feature projection

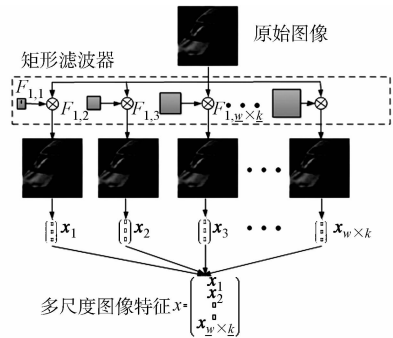


图 3 图像特征的多尺度表示
Fig. 3 Representation of multi-scale feature

知识, 在粒子群中, 粒子的行为可表示为

$$\mathbf{v}_i^{n+1} = \omega \mathbf{v}_i^n + \varphi_1 r_1 (\mathbf{x}_i^b - \mathbf{x}_i^n) + \varphi_2 r_2 (\mathbf{g} - \mathbf{x}_i^n), \quad (3)$$

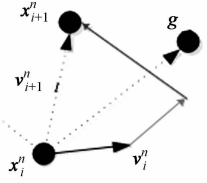
$$\mathbf{x}_i^{n+1} = \mathbf{x}_i^n + \mathbf{v}_i^{n+1}. \quad (4)$$

式(3), (4)中: ω 为迭代权重; φ_1, φ_2 为大于 0 的常数, 用于平衡个体最佳和全局最佳; $r_1, r_2 \in (0, 1)$ 为均匀分布的随机数。

在 PSO 算法中, 粒子迭代中都有一个对应的适应函数 $f(\mathbf{x}_i^{(t)})$, 当粒子在当前状态下的适应函数值大于个体最佳 \mathbf{x}_i^b 或全局最佳 \mathbf{g} 的适应函数值, 将当前粒子状态更新为个体最佳或全局最佳 \mathbf{g} . 如此反复迭代, 直到设定的迭代次数或搜索到最佳为止。

图 4 粒子更新示意图

Fig. 4 Particle update schematic



2.2 结合 PSO 改进的压缩跟踪

采用 PSO 方法对压缩跟踪搜索问题进行优化求解. 在当前这一帧搜索目标时, 在以前一帧目标位置为中心的边长为 R 的方形区域, 在多个尺度空间内, 先随机分布多个粒子, 提取这些粒子的 Haar-like 特征. 然后, 计算粒子的 Fitness, 更新和迭代粒子位置和速度, 直到收敛, 得出这一帧的目标位置. 最后, 在目标中心位置周围采集正样本和负样本, 更新朴素贝叶斯分类器. 收敛条件为: 当所有粒子的速度都为 0; 或达到最大迭代次数; 或满足匹配的最佳适应值已经过了 3 次迭代. 粒子 $\mathbf{x}_i^t = \{\mathbf{X}_i^t, \mathbf{Y}_i^t, \mathbf{w}_i^t, \mathbf{h}_i^t\}$ 包含位置和尺度空间. 种群中粒子个数为 K . 假设 t 时刻个体最佳集合为 $\{\mathbf{p}^{t,j}\}_{j=1}^K$, 为了保证粒子的多样性, 下一帧的初始都是在当前帧相应区域内高斯随机重分布^[12]. 其高斯分布为

$$\mathbf{x}_{j,0}^{t+1} \sim N(\mathbf{p}^{t,j}, \Sigma). \quad (5)$$

式(5)中: Σ 为高斯矩阵的协方差.

对角线元素为速度预测, 即

$$\mathbf{v}_{\text{pred}}^t = \mathbf{g}_{t-1} - \mathbf{g}_{t-2}. \quad (6)$$

粒子的速度则满足均匀分布 $\mathbf{x}_{j,0}^{t+1} \sim u(0, \mathbf{v}_{\text{pred}}^t)$.

2.3 适应值函数

在当前帧, 以前一帧为中心, 半径 R (足够大) 的圆, 或者为边长的正方形范围内, 按高斯分布随机初始化 K 个粒子. 由于在前一帧中已经更新了分类器, 因此, 用分类器的分类响应作为粒子 j 在第 t 次迭代时相应的适应函数值, 即

$$f(\mathbf{x}^{j,t}) = H(\mathbf{z}^{j,t}) = \log \left[\frac{\prod_{i=1}^n p(\mathbf{z}_i^{j,t} | y = 1) p(y = 1)}{\prod_{i=1}^n p(\mathbf{z}_i^{j,t} | y = 0) p(y = 0)} \right] = \sum_{i=1}^n \log \left(\frac{p(\mathbf{z}_i^{j,t} | y = 1)}{p(\mathbf{z}_i^{j,t} | y = 0)} \right). \quad (7)$$

式(7)中: $\mathbf{z}^{j,t}$ 为粒子 $\mathbf{x}^{j,t}$ 在当前位置时提取的低维特征向量.

根据每个粒子的适应函数值 $f(\mathbf{x}^{j,t})$, 得到全局最佳 \mathbf{g} 和粒子 j 的个体最佳 $\mathbf{p}^{t,j}$ 的更新, 即

$$\mathbf{p}^j = \begin{cases} \mathbf{x}^{j,t+1}, & f(\mathbf{x}^{j,t+1}) > f(\mathbf{p}^j), \\ \mathbf{p}^j, & \text{其他}, \end{cases} \quad (8)$$

$$\mathbf{g} = \arg \max_p f(\mathbf{p}^j). \quad (9)$$

为了防止粒子群的无限扩散与发散, 对加速参数 φ_1, φ_2 及最大速度 $\mathbf{v}_{\text{max}}^t$ 进行约束, 其参数满足

$$\varphi_1 = 2f(\mathbf{p}^j) / (f(\mathbf{p}^j) + f(\mathbf{g})), \quad (10)$$

$$\varphi_2 = 2f(\mathbf{g}) / (f(\mathbf{p}^j) + f(\mathbf{g})), \quad (11)$$

$$\mathbf{v}_{\text{max}}^t = 1.2 \times \mathbf{v}_{\text{pred}}^t. \quad (12)$$

由式(10)~(12)可知: 通过适应函数值可以控制个体和社会认知的变化, 使之得到平稳控制, 而粒子最大速度可利用先验知识获取.

2.4 分类器更新

当获取当前帧的最大响应位置时, 确定目标在当前帧的位置. 为了确定正负样本的采样范围, 通过 2 个距离阈值 α 和 β 来界定采样的类别. 在 $\mathbf{D}^\alpha = \{\mathbf{z} | \|\mathbf{l}(\mathbf{z}) - \mathbf{l}_t\| < \alpha\}$ 区域内采集正样本, 在区域 $\mathbf{D}^{\zeta,\beta} =$

$\{\mathbf{z}|\zeta<||\mathbf{l}(\mathbf{z})-\mathbf{l}_i||<\beta\}$ 采集负样本. 其中, $\alpha<\zeta<\beta$, 即靠近当前目标中心的随机采集为正样本, 外围采集为负样本.

根据正负样本, 分类器的参数更新为

$$\left. \begin{aligned} \boldsymbol{\mu}_i^1 &\leftarrow \lambda \boldsymbol{\mu}_i^1 + (1-\lambda) \boldsymbol{\mu}^1, \\ \boldsymbol{\delta}_i^1 &\leftarrow \sqrt{\lambda(\boldsymbol{\delta}_i^1)^2 + (1-\lambda)(\boldsymbol{\delta}^1)^2 + \lambda(1-\lambda)(\boldsymbol{\mu}_i^1 + \boldsymbol{\mu}^1)^2}. \end{aligned} \right\} \tag{13}$$

式(13)中: $\boldsymbol{\mu}_i^1, \boldsymbol{\delta}_i^1$ 分别为第 i 个特征正样本的均值和标准差; $\lambda>0$ 为分类器学习参数; $\boldsymbol{\mu}_i^1$ 和 $\boldsymbol{\delta}_i^1$ 的定义为

$$\left. \begin{aligned} \boldsymbol{\mu}_i^1 &= \frac{1}{n} \sum_{k=0|y=1}^{n-1} \mathbf{v}_i(k), \\ \boldsymbol{\delta}_i^1 &= \sqrt{\frac{1}{n} \sum_{k=0|y=1}^{n-1} (\mathbf{v}_i(k) - \boldsymbol{\mu}^1)^2}. \end{aligned} \right\} \tag{14}$$

负样本的均值和方差参数 $\boldsymbol{\mu}_i^0$ 和 $\boldsymbol{\delta}_i^0$ 的更新也与式(13)类似.

当目标在场景中因转向、形变、遮挡及镜头晃动等因素发生剧烈变化时, 可能导致粒子在给定范围内的适应值无法小于阈值, 搜索终止于局部最优. 结合文献[12]的方法, 采用选择更新和 re-search 策略. 当目标发生遮挡时, 分类器的响应值显著减小, 为了不被遮挡物完全取代, 设定一个分类阈值 T_c ; 当分类相应低于阈值时, 分类器停止更新; 当在给定的局部范围内无法搜索到目标时, 则跳出局部搜索范围, 增加粒子数量, 将 PSO 搜索范围扩大至全局搜索.

PSO-CT 算法有以下 5 个步骤.

步骤 1 初始化粒子群参数、分类器和目标特征.

步骤 2 对于每一帧, 有如下 5 点:

1) 计算 $\text{Frame}(t)$ 的积分图;

2) 如果第 $t-1$ 帧的分类响应 $P>T_c$, 则对于每个目标, 根据上一帧得到的目标位置, 在以上一帧的目标为中心, d 为边长的正方形内随机产生 K 个粒子, 并初始化个体最佳 \mathbf{p}^j 和全局最佳 \mathbf{g} ; 否则, 在全局范围内, 将尺度参数 w, h 范围扩大, 随机产生 $4K$ 个粒子, 并初始化 $4K$ 个例子的个体最佳 \mathbf{p}^j 和全局最佳 \mathbf{g} ;

3) 计算每个粒子的 Haar-like 特征, 计算粒子的分类得分, 得到各自的适应值 $f(\mathbf{x}^{(j,t)})$;

4) 比较 $f(\mathbf{x}^{(j,t)})$ 及 \mathbf{p}^j 和 \mathbf{g} , 更新个体最佳 \mathbf{p}^j 和全局最佳 \mathbf{g} , 修改粒子的位置和速度;

5) 如果 $f(\mathbf{g})$ 满足搜索结束条件, 则得到这一帧的目标位置: $\text{Loc}=\mathbf{g}$; 否则, 跳转至全局搜索 2).

步骤 3 在跟踪到的当前帧目标周围 $\mathbf{X}^{\zeta,\beta}=\{\mathbf{x}|\zeta<||\mathbf{I}_{t+1}(\mathbf{x})-\mathbf{I}_t(\mathbf{x}^*)||<\beta\}$, 采集正、负样本.

步骤 4 利用上面的积分图, 计算正样本和负样本的 Haar-like 特征值, 压缩特征.

步骤 5 通过正样本和负样本更新分类器, 进行下一帧处理.

3 结果与分析

3.1 粒子数的确定

为了验证算法的有效性, 在 CPU 主频为 3 GHz, 内存为 8 GB 的 PC 实验环境下, 采用几个标准视频集^[13], 将文中算法与粒子滤波(PF)、基于检测学习跟踪算法(TLD)、压缩跟踪算法(CT)进行比较.

为了更好确定实验中粒子个数, 针对粒子数与算法性能的关系进行多次重复实验, 结果如图 5 所示.

图 5 中: n_1 为粒子数; t 为帧处理时间. 由图 5 可知: 如要实现实时跟踪, 必须合理控制粒子数量. 其中, 中心定位误差(CLE)的计算为

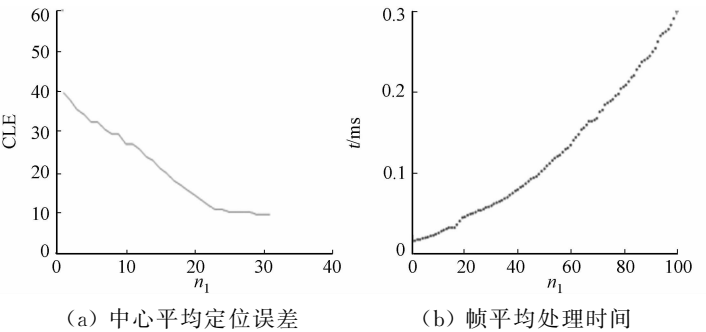


图 5 粒子数与平均定位误差和帧平均处理时间

Fig. 5 Particle number, CLE and PT

$$CLE = \sqrt{(x_t - x_r)^2 + (y_t - y_r)^2}.$$

(15)

式(15)中: (x_t, y_t) 和 (x_r, y_r) 分别表示当前跟踪定位结果与人工定位标准位置.

3.2 PSO-CT 的跟踪性能分析

将粒子数设置为 30, 将收敛阈值 $T_h=0.01, T_c=0.6$, 迭代参数 $\omega=0.5$, 初始 $\varphi_1=0.5, \varphi_2=0.5$, 最大迭代次数设为 100. 设 4 维搜索空间半径为 40. 在测试粒子搜索目标过程中, 计算复杂度的实验结果, 如图 6 所示. PSO-CT 平均每个粒子的迭代次数(n_2)为 42.4 次·帧⁻¹, 总迭代次数为 30×42.4, 相当于只需 1 272 次搜索和比对, 就能在每一帧图像中发现目标. 文献[6]采用改进全搜索的 coarse-to-fine 策略需要 3 496 次. 在几组复杂场景下的跟踪结果, 如图 7, 8 所示, 其帧率如表 1 所示. 图 8 中: n_3 为帧数; e 为跟踪误差.

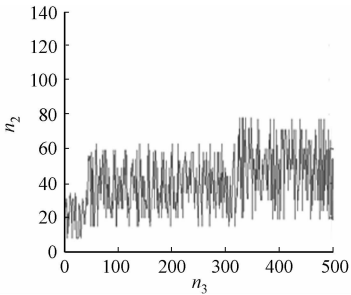


图 6 粒子平均迭代次数
Fig. 6 Average iteration times of particle

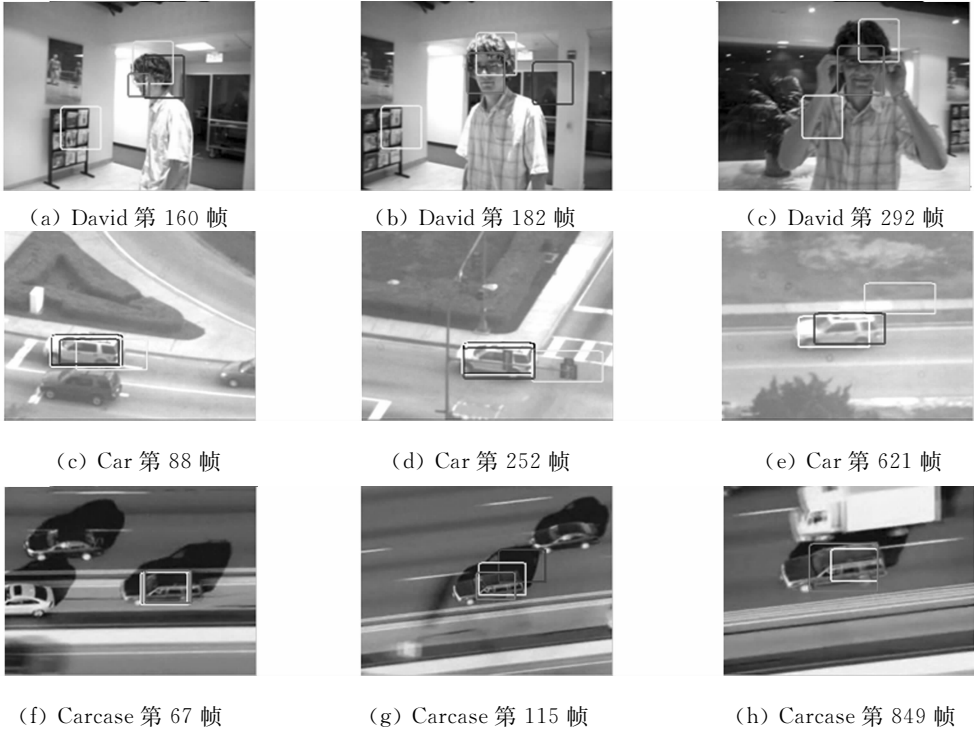


图 7 复杂环境下几种跟踪算法结果

Fig. 7 Tracking result of some algorithm in complex scenes

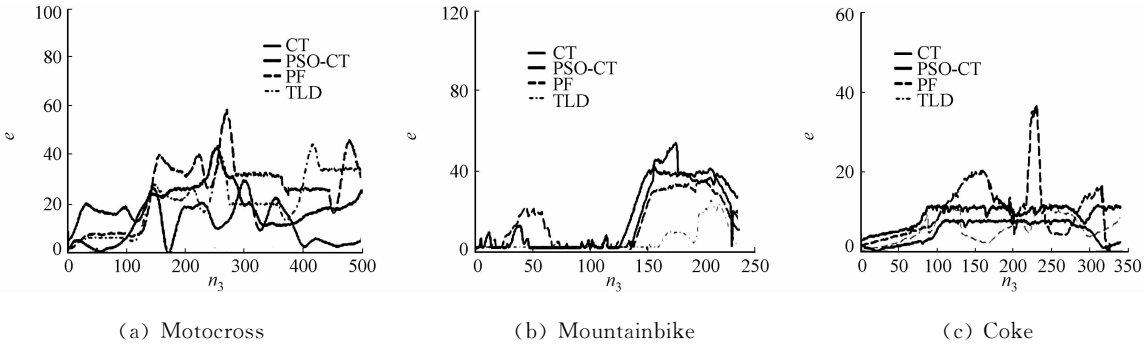


图 8 几种算法的跟踪误差

Fig. 8 Tracking error of some algorithms

由图 8 可知: Motocross 中目标因为翻转产生平稳形变, 此时, TLD 算法、PF 算法和 PSO-CT 算法表现出良好效果; Mountainbike 中场景从平稳跨越到剧烈抖动, 此时, 原始 PSO-CT 算法在平稳环境误差较小, 而面对突变, 最终也能快速校正偏差; Coke 场景中, 中间经历几次遮挡和高度变化, 当镜头或者

景深发生明显变化时,多尺度的 PSO-CT 的尺度适应性能优势明显.

表 1 几种跟踪算法的平均帧率
Tab.1 Average frame processing rate of some tracking algorithms

数据集	分辨率/px×px	平均帧率/帧·s ⁻¹			
		PSO-CT	CT	PF	TLD
Motocross	470×310	56.4	54.5	29.9	47.1
Mountainbike	352×288	63.3	59.2	36.5	56.4
David	320×240	91.3	86.8	40.2	66.0
Carcase	291×218	104.2	99.8	53.3	75.6

综上可知,在变化比较快的复杂场景下,CT 和 PSO-CT 表现相对稳定,且 PSO-CT 能适应突然、快速的目标位置变化.当镜头或者景深发生明显变化时,多尺度的 PSO-CT 在尺度变化的自适应方面体现出良好的跟踪效果.在时间效率上,如表 1 所示.由于采用特征压缩方法,将特征用低维的向量表示,CT 算法和 PSO-CT 算法在速度上具有优势,体现出良好的快速性和实时性.

4 结束语

针对原始压缩跟踪存在的不足,结合粒子群优化算法的优化多维空间搜索,实现复杂场景下多尺度目标跟踪.实验表明,提出的 PSO-CT 跟踪方法在实时性和稳定性上具有一定优势.该算法对运动目标发生部分遮挡和形变时仍然有较好的鲁棒性,然而,当多个目标之间产生聚合与分离时,跟踪还达不到最理想的结果,这将是下一步工作需要考虑和突破的地方.

参考文献:

[1] 龚声蓉,刘纯平,季怡.复杂场景下图像与视频分析[M].北京:人民邮电出版社,2013:8-9.

[2] KALAL Z,MIKOLAJCZYK K,MATAS J. Tracking-learning-detection[J]. Pattern Analysis and Machine Intelligence,2012,34(7):1409-1422.

[3] STALDER S,GRABNER H,GOOL L V. Cascaded confidence filtering for improved tracking-by-detection[M]. Berlin:Springer,2010:369-382.

[4] GRABNER H,GRABNER M,BISCHOF H. Real-time tracking via on-line boosting[C]// British Machine Vision Conference. Edinburgh:[s. n.],2006:6.

[5] FELZENSZWALB P,MCALLESTER D,RAMANAN D. A discriminatively trained, multiscale, deformable part model[C]// IEEE Conference on Computer Vision and Pattern Recognition. Anchorage:IEEE Press,2008:1-8.

[6] ZHANG Kaihua,ZHANG Lei,YANG M H. Real-time compressivetracking[C]// Europeanence on Computer Vision. Berlin:Springer,2012:864-877.

[7] SAINI S,AWANG R D R,ZAKARIA M N B,*et al.* A review on particle swarm optimization algorithm and its variants to human motion tracking[J]. Mathematical Problems in Engineering,2014(2014):1-16.

[8] ZHANG Xiaoqin, HU Weiming, MAYBANK S, *et al.* Sequential particle swarm optimization for visual tracking [C]//IEEE Conference on Computer Vision and Pattern Recognition. Anchorage:IEEE Press,2008:921-928.

[9] FLEISCHMANN P,AUSTVOLL I,KWOLEK B. Particle swarm optimization with soft search space partitioning for video-based markerless pose tracking[J]. Lecture Notes in Computer Science,2012,7517:479-490.

[10] BARANIUK R. Compressive sensing[J]. IEEE signal processing magazine,2007,24(4):118-121.

[11] THIDA M,ENG H,MONEKOSSO D N,*et al.* A particle swarm optimisation algorithm with interactive swarms for tracking multiple targets[J]. Applied Soft Computing,2013,13(6):3106-3117.

[12] TAWAB A M A,ABDELHALIM M B,HABIB S. Efficient multi-feature PSO for fast gray level object-tracking [J]. Applied Soft Computing,2014,14(1):317-337.

[13] WU Yi,LIM J,YANG M H. Online object tracking: A benchmark[C]// IEEE Conference on Computer vision and pattern recognition. Portland:IEEE Press,2013:2411-2418.

(责任编辑:钱筠 英文审校:吴逢铁)