

文章编号 1004-924X(2012)11-2540-08

结合标准对冲与核函数稀疏分类的目标跟踪

匡金骏*, 柴毅, 熊庆宇
(重庆大学 自动化学院, 重庆 400030)

摘要:针对经典稀疏分类目标跟踪算法在噪声, 遮挡等恶劣环境下精度不高的问题, 提出了一种新的目标跟踪算法。该算法在标准对冲框架下结合了核函数稀疏分类方法以及自适应字典更新方法, 能够较好地适应类间相似度较高与目标外形变化较大等恶劣情况。核函数技巧能够增强分类器性能, 但通用方法求解凸优化问题的效率较低, 不能满足目标跟踪问题的实时性要求, 故提出用核函数随机坐标下降(KRCD)算法来高效求解稀疏系数, 并使用核函数稀疏分类方法(KRCD-SRC)来计算各个粒子的代价值。为了避免模板漂移问题, 解释了目标字典和背景字典的在线更新方法。最后, 结合标准对冲算法估算目标的状态信息。在使用 50 个粒子进行跟踪时, 本文算法的处理帧率能够达到 14 frame/s。相比其它几种经典目标跟踪算法, 本文算法具有更好的精确性和鲁棒性。

关键词:目标跟踪; 核函数稀疏分类; 自适应字典更新; 标准对冲

中图分类号: TP391 **文献标识码:** A **doi:** 10.3788/OPE.20122011.2540

Visual object tracking combined normal hedge and kernel sparse representation classification

KUANG Jin-jun*, CHAI Yi, XIONG Qing-yu

(College of Automation, Chongqing University, Chongqing 400030, China)

* Corresponding author, E-mail: kuangjinjun@gmail.com

Abstract: To achieve the robust tracking for a visual object under challenging conditions in the noisy, occlusion and the deformation, a novel visual object tracking method is proposed in this paper. By combining the Kernel Sparse Representation Classification (KSRC) and adaptive dictionary updating method under Normal Hedge framework, this method can handle tough situations like high inter-class similarities and drastically target appearance variations. Although the KSRC enhances classification performance, standard convex optimization is not fast enough for tracking in real time. Thus an efficient Kernel Random Coordinate Descent(KRCD) method is proposed to calculate the sparse coefficient vector, and the KRCD-SRC classification method is taken to calculate the loss value of each particle. In order to avoid the template drifting, the adaptive dictionary updating method is also given. At last, the states of the target are estimated by the Normal Hedge. Experiments show that the average computing frame rate of the proposed method is 14 frame/s when 50 particles are used. Extensive test results suggest that the proposed method outperforms several state-of-art tracking methods in

收稿日期: 2012-05-04; 修订日期: 2012-07-02.

基金项目: 国家自然科学基金资助项目(No. 60974090); 国家教育部博士点基金资助项目(No. 102063720090013); 中央高校基本科研业务费资助项目(No. GDJXS10170010)

many complex conditions.

Key words: object tracking; kernel sparse representation; adaptive dictionary updating; Normal Hedge

1 引 言

视频目标检测与跟踪问题是机器视觉的基础问题之一,它是目标识别、行为识别等后续应用的基础,在社会安全、航空航天等重大项目中有着广阔的应用前景。多年来,众多学者就目标跟踪问题提出了多种方法,包括基于轮廓的方法^[1]、基于核的方法^[2]、基于统计的方法^[3]以及基于稀疏分类的方法^[4-6]等。其中,基于稀疏分类框架的目标跟踪算法因其独特的抗噪声与遮挡能力在近期引起了较多学者的关注。Zhenjun Han 等在文献[4]中结合稀疏分类与卡尔曼滤波(Kalman Filtering)方法,提出了 AdaSR 算法,该算法能够保证目标模板的时域连续性,对目标模板的形变也不敏感。Xue Mei 等在文献[5]中提出了结合稀疏分类与贝叶斯推理框架的目标跟踪与识别算法,在此文献中,作者还提出了用细节模板对噪声,遮挡等干扰进行建模的方法,并提出了一种非负约束用以增强稀疏分类精度。Qing Wang 等在文献[6]中提出了一种基于判别方法的稀疏分类目标跟踪算法。此算法的特点是能够从原始图像中学习具有判别性的字典用以增强跟踪的精度。

虽然稀疏分类方法在目标跟踪问题中得到了较多的应用,但近期文献[7-8]指出,对于小样本分类问题,系数的稀疏性对分类准确率并没有实质的帮助。针对此问题,Ke Huang 等在文献[9]中指出结合线性判别分析技术能够提升类间的区分度,提升稀疏分类效果。Shenghua Gao 等在文献[10]中成功将核函数(Kernel)技巧与稀疏分类结合在了一起,提出了基于 Feature-Sign Search (FSS)的核函数稀疏分类(KSRC)算法并将其成功应用于人脸识别中。然而,Cuicui Kang 等在文献[11]中指出使用 FSS 方法求取 KSRC 中凸优化问题的效率较低,故提出了核函数坐标下降法(KCD)用以求解凸优化问题,并结合 LBP 特征构建了人脸识别系统。但是目标跟踪问题的实时性

要求比人脸识别系统要高,针对此问题,本文提出了核函数随机坐标下降法(KRCD)用以求解 KSRC 框架中的凸优化问题。

目标跟踪是一种 2 类在线分类问题,其需要一套特殊的机制用以完成样本的抽样工作。近年来,粒子滤波(蒙特卡洛)方法因其能够较完美地解决此问题而受到众多学者的关注。龚俊亮等在文献[12]中提出了一种结合粒子滤波和马尔科夫随机场的目标跟踪算法,用于解决多目标跟踪时粒子数量呈指数级增长的问题。吴川等在文献[13]中提出了一种基于粒子滤波的跟踪和检测合一的系统,并证明了系统具有一定的稳定性。然而,退化(Degeneracy)现象的发生是粒子滤波的主要缺陷^[14]之一,随着滤波迭代次数的增加,大部分粒子的权重会变得很小,而只有很少的粒子具有较大的权重。针对以上问题,Kamalika 等^[15]在近期提出了一种类似于粒子滤波的在线机器学习算法:标准对冲 Normal Hedge(NH),其定义了一种独特的粒子重采样方法能够有效地避免粒子退化问题。另外,在此文献中还证明了粒子集合整体的懊恼值上限,即粒子集合的懊恼值接近所有粒子中最好的一群,所以标准对冲算法在即使只有很少一部分粒子状态接近目标的真实状态,算法也能较准确地估计出目标的真实状态。此算法在近期在声源定位^[16]等应用中有较好的表现。

本文提出的目标跟踪算法能够有效解决以上问题,其特点在于:

(1)提出了核函数随机坐标下降法(KRCD)用以求解凸优化问题,其效率比 KCD 更高;同时,也提出了核函数稀疏分类 KRCD-SRC 方法,其能够有效求取粒子的代价值。

(2)自适应字典更新:给出了通用更新方式与时序循环更新方式,并通过实验证明其能在目标形变,被遮挡等恶劣情况下提升系统的稳定性。

(3)标准对冲框架:将 KRCD-SRC 方法嵌入 NH 框架内作为各个粒子的代价函数,在相同的跟踪精确度前提下,避免了经典粒子滤波的粒子

退化问题,减少了所需粒子数量,减少了系统的计算量。

与现存基于稀疏表示框架的图像跟踪算法相比,本算法在目标外形变化、遮挡、光照改变、背景与目标存在相似特征等复杂情况时具有更稳定的跟踪效果。

2 标准对冲

标准对冲算法使用一组粒子对目标的可能状态(位置、面积、旋转等)进行采样并计算其与目标模板之间的差异(代价值 $l_{i,t}$),然后通过迭代运算估算目标的隐藏态。标准对冲算法算法:

首先初始化 $N \geq 2$ 个粒子。对于粒子集,有对应于每个粒子的权值 $\omega_i = (\omega_{i,1}, \dots, \omega_{i,n})$ 和与每个粒子所对应的代价值 $l_{i,t}$,在第 t 轮迭代整体代价值为:

$$l_{\Lambda,t} = \sum_{i=1}^N \omega_i l_{i,t}, \quad (1)$$

其中: $l_{i,t} = \delta(i)$ 。

粒子 i 在时刻 t 的懊恼值为 $r_{i,t} = l_{\Lambda,t} - l_{i,t}$;在时刻 t 的累积懊恼值为

$$R_{i,t} = r_{i,t} + (1 - \lambda) \cdot R_{i,t-1}. \quad (2)$$

如果 $R_{i,t} < 0$ 则在上时刻目标位置以高斯分布重置此粒子。(此步骤避免了粒子退化问题)

标准对冲基于半正态分布函数,为:

$$\Phi(x, c) = \exp\left\{-\frac{(\max(0, x))^2}{2c}\right\}, \quad (3)$$

其中: $x \in \mathbf{R}, c > 0$ 。除了懊恼值 $R_{i,t}$,标准对冲算法还包含一个尺度参数 c_i 。 $R_{i,t}$ 与 c_t 满足下式:

$$e = \frac{1}{N} \sum_{i=1}^N \exp\left\{-\frac{(\max(0, R_{i,t}))^2}{2c_t}\right\}. \quad (4)$$

在时刻 t 分配给粒子 i 的权值为:

$$\omega_{i,t} \propto \frac{\partial}{\partial x} \varphi(x, c) \Big|_{x=R_{i,t-1}, c=c_{t-1}} = \frac{[R_{i,t-1}]_+}{c_{t-1}} \exp\left\{-\frac{([R_{i,t-1}]_+)^2}{2c_{t-1}}\right\}. \quad (5)$$

目标状态为:

$$x_{\Lambda,t} := \sum_{i=1}^N \omega_{i,t} \cdot x_{i,t}. \quad (6)$$

文献[15]证明了标准对冲算法的整体懊恼值上限为 $O(\sqrt{T \ln N} + \ln^2 N)$,其中 N 为粒子数量,即粒子集合的懊恼值接近所有粒子中最好的一群。图 1 显示了将标准对冲算法应用于视频跟踪

实验中各粒子所对应的权值。由图可知,在 1~300 帧左右,各粒子的权值分布均匀,这说明跟踪过程较为平稳,但 305 帧左右跟踪过程受到干扰,大多数粒子对应的权值下降,即使只有少量粒子成功跟踪目标也能保证整个跟踪过程的稳定性。

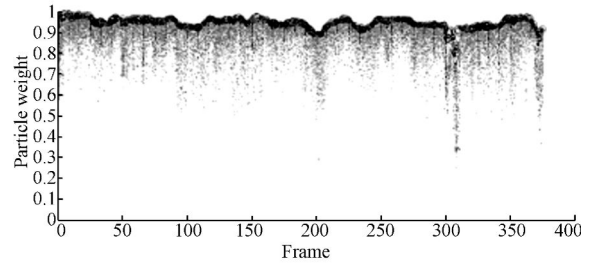


图 1 标准对冲算法粒子权值分布

Fig. 1 Particle weight distribution of normal hedge method

3 核函数稀疏分类算法

目标跟踪问题可考虑为基于目标和背景的两类在线分类问题,本节介绍稀疏分类的原理以及核函数稀疏分类的原理。最后将介绍粒子代价值 $l_{i,t}$ 的求法。

3.1 稀疏分类

设目标和背景图的分辨率为 $w \times h$,将每幅图变换为一个列向量 $\mathbf{v} \in \mathbf{R}^d, d = w \times h$,那么 n 个训练图像组成字典矩阵(也称基矩阵) $\mathbf{D} = [\mathbf{v}_1, \dots, \mathbf{v}_n] \in \mathbf{R}^{d \times n}$,令 $\mathbf{D}_1 \in \mathbf{R}^{d \times k}$ 表示目标字典, $\mathbf{D}_2 \in \mathbf{R}^{d \times (n-k)}$ 表示背景字典,则 $\mathbf{D} = [\mathbf{D}_1, \mathbf{D}_2]$ 。

在稀疏表示框架下,纯净的目标列向量 $\mathbf{y} \in \mathbf{R}^d$ 可由训练矩阵 \mathbf{D} 和稀疏系数向量 \mathbf{C} 的线性组合来进行重构^[4]:

$$\mathbf{y} = \sum_{i=1}^n \mathbf{v}_i \cdot c_i, \quad (7)$$

其中:稀疏系数 $\mathbf{C} = [0, \dots, 0, c'_1, 0, \dots, c'_l, \dots, 0]^T \in \mathbf{R}^n$,在理想情况下,根据字典的构造,系数向量中只有数个系数不为零。

在现实情况中,因为目标列向量中有噪声信号污染,目标列向量很难得到完美恢复;所以(7)可等价写为:

$$\mathbf{y}_0 = \mathbf{y} + \mathbf{e} + \sum_{i=1}^n \mathbf{v}_i \cdot c_i + \mathbf{e}, \quad (8)$$

其中: \mathbf{y}_0 为实际列向量, $\mathbf{e} \in \mathbf{R}^d$ 为目标的真实列向

量与重构列向量之差。在式(8)中, \mathbf{v}_i 与 \mathbf{y}_0 是已知的, 需要求出的是测试图像在字典中所对应的系数向量 \mathbf{C} 。因为 \mathbf{D} 是欠定的, 所以没有唯一解。式(8)的最稀疏解可表示为:

$$\min_{\mathbf{C}} \left[\left\| \mathbf{y}_0 - \sum_{i=1:n} \mathbf{v}_i \cdot c_i \right\|_2^2 + \lambda \|\mathbf{C}\|_0 \right], \quad (9)$$

其中: l_0 范数 $\|\mathbf{C}\|_0$ 计算非零系数的个数, λ 为折中权值, λ 增大则系数的稀疏度增大。式(9)为一个无法在多项式时间中求解的 NP-hard 问题, 所幸近期压缩感知理论^[17]证明如果式(9)的解足够稀疏, 则它等于:

$$\min_{\mathbf{C}} \left[\left\| \mathbf{y}_0 - \sum_{i=1:n} \mathbf{v}_i \cdot c_i \right\|_2^2 + \lambda \|\mathbf{C}\|_1 \right], \quad (10)$$

即用 l_1 范数方法代替 l_0 范数。此为一个凸优化问题, 存在多种有效的解法。

如果根据以上计算得到的系数 $\mathbf{C} \in \mathbf{R}^n$ 足够稀疏, 则其包含了有效的分类信息。具体来说, 将背景字典所对应的系数置零, 目标残差 $r_1 = \left\| \mathbf{y}_0 - \sum_{i=1:k} \mathbf{v}_i \cdot c_i \right\|_2^2$, 其中原子 1 到 k 属于目标字典。同理, 将目标字典所对应的系数置零, 背景残差 $r_2 = \left\| \mathbf{y}_0 - \sum_{i=k:n} \mathbf{v}_i \cdot c_i \right\|_2^2$, 其中原子 k 到 n 属于背景字典。则

$$\begin{cases} \text{label}_y = 1, & \text{若 } r_1 < r_2 \\ \text{label}_y = 0, & \text{若 } r_2 < r_1 \end{cases}, \quad (11)$$

其中: label_y 为图像 \mathbf{y}_0 所对应的类别。

3.2 核函数稀疏分类

$\mathbf{D} = [\mathbf{v}_1, \dots, \mathbf{v}_n]$ 的列向量之间不相关。但是在跟踪任务中, 有时目标和背景之间有较多的相似特征, $\mathbf{D}_1, \mathbf{D}_2$ 的列向量的相关度高, 则式(10)的非零解可能集中在错误的类别上^[8]。这种现象的另外一种解释为, 目标字典 \mathbf{D}_1 与背景字典 \mathbf{D}_2 间区分度不高, 这将影响跟踪的精确度。

核函数(Kernel)为一种维度变换函数, 能将特征向量由低维空间投影到高维度空间, 增大类别间的几何距离, 使得线性不可分变为线性可分。其在支持向量机(Support Vector Machine)等经典分类器中有较多的成功应用。受文献^[10]的启发, 核函数技巧也能应用于稀疏表示问题中, 降低分类难度, 减少重构误差。

设函数 $\varphi(\cdot)$ 将特征从维度空间 d 投影到维度空间 $h: \mathbf{R}^d \rightarrow \mathbf{R}^h$, ($d < h$), 待分类图像向量被投影为 $\mathbf{y} \rightarrow \varphi(\mathbf{y})$, 字典被投影为 $\mathbf{D} \rightarrow \tilde{\mathbf{D}} = (\varphi(\mathbf{v}_1), \varphi(\mathbf{v}_2), \dots, \varphi(\mathbf{v}_n))$, 设 $K(\mathbf{x}, \mathbf{x}) = \varphi(\mathbf{x})^T \varphi(\mathbf{x}) = 1$ 。则

式(11)在核函数空间中可写为

$$\min_{\mathbf{C}} J(\mathbf{C}) = \left[\left\| \sum_{i=1:n} \varphi(\mathbf{v}_i) \cdot c_i - \varphi(\mathbf{y}_0) \right\|_2^2 + \lambda \|\mathbf{C}\|_1 \right]. \quad (12)$$

受文献^[18]的启示, 本文提出核函数随机坐标下降法(Kernel Random Coordinate Descent, KRCD)用来解式(12), 与核坐标下降法(KCD)^[11]以及 feature-sign search^[10]方法相比, 本算法具有更快的收敛速度。当 $c_i \neq 0$ 时, 对 $J(\mathbf{C})$ 求偏导, 则

$$\nabla J(\mathbf{C}) = \sum_{i=1:n} c_i \cdot K(\mathbf{v}_j, \mathbf{v}_i) - K(\mathbf{v}_j, \mathbf{y}) + \lambda \text{sign}(c_j), \quad (13)$$

其中: $k(\mathbf{v}_i, \mathbf{v}_i) = \varphi(\mathbf{v}_i)^T \cdot \varphi(\mathbf{v}_i)$, $k(\mathbf{v}_i, \mathbf{y}) = \varphi(\mathbf{v}_i)^T \cdot \varphi(\mathbf{y})$ 。由文献^[16]可知, 在时刻 $k+1$ 可使用式(13)更新 c_i 的条件为: $J(\mathbf{C})$ 的梯度为独立坐标利普希茨连续(Coordinate-wise Lipschitz Continuity)

$$c_i^{k+1} = c_i^k - L_i^{-1} \cdot \nabla J(\mathbf{C}) \cdot \mathbf{e}_i, \quad (14)$$

其中: i 为 $[0, n]$ 之间的随机整数, L_i 为 $\nabla J(\mathbf{C})$ 所对应的利普希茨常数, $\mathbf{e}_i = [0, \dots, p, \dots, 0]$ 且 $p = 1$ 。 $\nabla J(\mathbf{C})$ 的梯度为独立坐标利普希茨连续的充分条件为 $|\nabla J_j(\mathbf{C}+h) + \nabla J_j(\mathbf{C})| \leq L_j |h|$, 通过计算可知 $L_j = K(\mathbf{v}_j, \mathbf{v}_j) = 1$, $|\nabla J_j(\mathbf{C}+h) - \nabla J_j(\mathbf{C})| = h$, 又因为 $|h| \in [0, 1]$, 故(13)条件成立。所以, 对于任意核函数 $K(\cdot, \cdot)$, KRCD 算法能够靠式(13)与(14)迭代优化稀疏系数 \mathbf{C} 。

由文献^[19]可知, KRCD 算法的复杂度为 $O\left(\frac{n}{\epsilon} \log \frac{1}{p}\right)$, 其中 ϵ 为重构精度, p 为信心水平(confidence level)且 $0 < p < 1$, 而 n 为字典中原子的数量。

本文提出 KRCD-SRC(Kernel Random Coordinate Descent based Sparse Representation Classification)方法用于对图像分类:

$$r_s = \left\| \sum_{i=1:s} \varphi(\mathbf{v}_i) \cdot c_i - \varphi(\mathbf{y}_0) \right\|_2^2 = (\mathbf{C}_s^T \mathbf{K}(\mathbf{v}_i, \mathbf{v}_i)_{n \times n} \mathbf{C}_s - 2\mathbf{K}(\mathbf{v}_i, \mathbf{y})_{n \times 1}^T \mathbf{C}_s), \quad (15)$$

其中: r_s 为对应于类 s 的残差, 1 到 s 为第 s 类对应的系数。对于 2 类分类问题, 则

$$\begin{cases} \text{label}_y = 1, & \text{若 } r_1 < r_2 \\ \text{label}_y = 0, & \text{若 } r_2 < r_1 \end{cases}, \quad (16)$$

其中: label_y 为图像向量 \mathbf{y}_0 所对应的类别。对于多类分类问题, 则分类结果为最小残差 r_s 对应的

类别。标准对冲算法使用的代价值 $l_{i,t} = r_t$ 。

本节提出的 KRCD 方法可结合多种核函数,如高斯核函数(Gaussian Kernel),多项式核函数(Polynomial Kernel),S 形核函数(Sigmoid Kernel)等。不同种类的核函数所对应的最终分类精度有所不同(如高斯核函数在小样本情况下分类精度较差),本文将在第 6 章显示对比实验结果。

4 自适应字典更新

文献[5]指出对于稀疏分类算法,使用原始图像作为字典,能够达到最好的分类效果,所以本文将图片直接转换为字典。为了保证字典能够表示目标的各种外观,必须要及时对字典进行更新。与只包含目标字典 D_1 的稀疏表示框架跟踪系统^[4-5]不同,本文算法不仅包含 D_1 ,而且在每个时刻都更新背景字典 D_2 。对背景建模能够有效增加系统字典的完备性,提高跟踪精度。设跟踪抽样图像与现存目标字典之间的残差为 r_t 。定义最小阈值 ξ_{\min} 与最大阈值 ξ_{\max} 。

如图 2 所示,如果 $r_t < \xi_{\min}$,则表示抽样图像 i 能够被目标字典表示,此抽样图像无需被转化为新的目标字典。如果 $r_t > \xi_{\min}$,则表示抽样图像 i 完全不能由目标字典表示,并且系统不能保证此抽样为目标。通用模板更新:如果 $\xi_{\min} < r_t < \xi_{\max}$,此抽样图像为目标模板,转化为目标字典。图 2 中,③、⑦、⑧和⑨都将被转化为目标模板。在大部分时候,通用模板更新方法结合时空约束能够保证目标字典是纯净的;但在目标外观变化程度既大速度也快的情况下,无法保证目标字典的完备性。对此问题,本文提出基于时序循环模板更新方法:设循环时间阈值为 θ ,考虑抽样图像 i 在时间 θ 内残差由 $r_t < \xi_{\max}$ 变为 $r_t > \xi_{\max}$,这说明抽样图像已经不再能由目标字典稀疏表示,但系统不能完全确定跟踪失败,此时将所有抽样图像保存在临时抽样池中,若在时间阈值 θ 内,残差由 $r_t > \xi_{\max}$ 变为 $r_t < \xi_{\max}$,说明跟踪没有失败,则临时抽样池中的所有模板都可作为候选目标字典。注意将候选目标字典转换为目标字典也需遵循稀疏原则,即按顺序检查候选目标字典 i 的残差,若 $\xi_{\min} < r_t < \xi_{\max}$ 则转换目标字典,反之则删除。如果在时间阈值 θ 内残差保持 $r_t > \xi_{\max}$,则说明跟踪失败。图 2 中,时序循环方法接受③~⑨。结合通

用与时序循环更新方法能够收集目标的各种外观,同时保证目标字典的稀疏性和完整性。背景字典 D_2 的创建方式较为简单,对于每帧图像,系统在目标周围进行随机采样,然后将样本转化为背景字典。

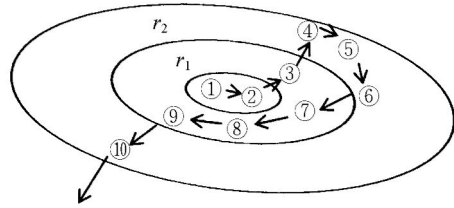


图 2 字典的在线更新

Fig. 2 Illustration of dictionary updating method

在长时间跟踪过程中,由选择机制创建的目标字典可能变得非常庞大,这将导致求解稀疏系数速度降低,制约跟踪性能。所以设时间阈值 ψ ,若目标字典 D_1^i 的存在时间超过阈值 $t > \psi$,则删除 D_1^i 以保证算法的收敛速度。最后,在每个时间点,系统删除并更新所有旧的背景字典。

5 算法实现

跟踪算法步骤如下:

步骤 1. 在视频某帧手动选取目标。并更新目标字典与背景字典。

步骤 2. 使用核函数稀疏分类方法计算各个粒子的代价值。

步骤 3. 使用标准对冲算法计算目标状态。

步骤 4. 更新目标字典与背景字典。

步骤 5. 读入下一帧图像并返回步骤 2。

6 实验与仿真

为了验证本文提出的目标跟踪算法的有效性和稳健性,本节进行两种类型的实验,第 1 类实验对比使用不同类别核函数在最优参数时的跟踪精度。第 2 类实验比较本文算法与其它 3 类经典跟踪算法跟踪精度。实验平台为基于 Windows 7 操作系统的 PC 计算机,处理器频率为 2.2 G,内存容量为 2 G,使用 C++ 编程。图像序列已经过预处理,分辨率为 640×480 。

6.1 多种核函数效果对比实验

对于核函数的选择问题,在现有文献中缺乏有效的指导原则,各种分类实验实验观察结果表明,特定问题使用特定的核函数会取得更好的效果。在众多核函数中,高斯核函数的应用范围最广泛,但其在训练样本数量较少时分类精度不佳。本实验将使用高斯核函数,多项式核函数以及 S 形核函数进行对比实验。

表 1 不同核函数的性能

Tab.1 Performance of different kernels

核函数	误差 (50 次)	方差 (50 次)	每帧耗时	综合评价
高斯核	9.67	3.61	0.061	5.22
多项式核	6.32	2.12	0.063	3.39
S 形核	7.21	3.05	0.069	3.93

高斯核函数(Gaussian Kernel)的数学表达式为:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left[-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{\sigma^2}\right]. \quad (17)$$

多项式核函数(Polynomial Kernel)的数达表达式为:

$$K(\mathbf{x}_i, \mathbf{x}_j) = (\mathbf{x}_i \cdot \mathbf{x}_j + n)^q, \quad (18)$$

S 形核函数(Sigmoid Kernel)的数学表达式为:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \tanh(v(\mathbf{x}_i \cdot \mathbf{x}_j) + t). \quad (19)$$

线性核函数不进行特征维度空间转换,若本文算法结合线性核函数,则等于稀疏表示分类算法 SRC^[5]。参数选择会影响其余核函数的分类精度,表 1 显示了 4 种算法在最优参数设定时的跟踪精度。其中综合评价指数定义与文献[20]相同,设待评价方法分别为 $\{X_1, X_2, \dots, X_n\}$, 评价指标分别为 $\{C_1, C_2, \dots, C_n\}$, 各评价指标的权重分别 $\{\omega_i\}_{i=1, m}$, $\sum \omega_i = 1$, $C_i(X_k)$ 是算法 X_k 对应于评价指标 C_i 的性能指数。算法 X_k 综合评价指数为

$$CEI(X_k) = \sum_{i=1}^m \frac{C_i(X_k)}{\max_{1 \leq j \leq n} C_i(X_j)} \times \omega_i, \quad (20)$$

式中 CEI 值越小,表明算法的综合性能越好。本实验设误差的权值为 0.5,方差权值为 0.1,耗时的权值为 0.4。本实验使用目标跟踪公用测试视频 PETS2009: S2-L2-Time-14-55-View_001。此视频的难点在于目标多次被遮挡,来往行人与目

标有很多相似的特征,并且目标的尺度变化较大。

表 1 中的“误差”代表算法求出的目标位置与手工标定的位置误差(单位为像素),误差,方差和每帧耗时结果为 50 次实验结果均值。在本实验中,高斯核函数的 σ^2 取值为 3.5,多项式核函数阶次 q 最优取值为 4,S 形核函数的参数最优取值为 $v=2, t=0.8$ 。粒子数为 50。由综合评价可知,多项式核函数在最优化参数情况下跟踪效果最佳。虽然每帧耗时有增加,但每秒帧率(FPS)均超过 14,实时性较强。

6.2 复杂情况下的跟踪效果对比实验

实验 2 的目的是对比本文算法,AdaSR^[4]以及粒子滤波(PF)^[13]在人工加噪,目标形变与被遮挡以及目标与背景有相似特征时的跟踪效果。评比的指标与 6.1 节相同,分别为误差,方差,每帧耗时以及综合评价。选择 AdaSR 和 PF 进行对比实验的原因是 AdaSR 算法结合了经典的稀疏分类 SRC 与卡尔曼滤波算法。PF 算法使用颜色特征以及 50 个粒子完成跟踪任务。如图 3 所示,本实验使用的视频分别为:snow(人工加 10 dB 高斯白噪声),PETS2006-S1-T1-C-1 测试集以及 PETS2007- background 测试集。本文算法使用 4 阶多项式核函数与 50 个粒子完成跟踪任务。为了保证结果的公正性,表 2~表 4 中误差,方差和每帧耗时为 50 次实验结果的平均值。



图 3 测试视频示例

Fig.3 Illustration of test videos

表 2 显示了在 PETS2007- background 测试集的测试结果,即目标和背景具有相似特征时 3 种算法的跟踪效果。实验结果显示本文算法的计算复杂度相比另外两种算法更多,但本文算法在综合得分中大幅度超越了另外两种算法,表明核函数稀疏分类方法具有提升跟踪精度的效果。

表 2 目标和背景具有相似特征的跟踪结果

Tab. 2 Tracking results for target with similar background patterns

算法	误差 (50 次)	方差 (50 次)	每帧耗时	综合评价
本文算法	3.33	2.89	0.062	1.97
AdaSR	12.67	4.61	0.045	6.81
PF	13.32	3.12	0.054	6.99

表 3 显示了 3 种算法在 PETS2006-S1-T1-C-1 视频中的测试结果。本次测试目标有形变和部分被遮挡。由结果可知,本文算法的自适应字典更新功能能够鲁棒地应对目标形变,且稀疏分类方法能够较好地处理部分遮挡问题。所以,本文算法与 AdaSR 算法在目标被部分遮挡时表现出了较稳定的跟踪结果。经典 PF 算法在本次测试中表现结果不佳。

表 3 目标形变与部分被遮挡的跟踪结果

Tab. 3 Tracking results for target with deformation and partial occlusion

算法	误差 (50 次)	方差 (50 次)	每帧耗时	综合评价
本文算法	3.61	2.68	0.061	2.09
AdaSR	9.42	1.32	0.047	4.86
PF	15.37	2.39	0.051	7.94

表 4 显示了 3 种算法在具有 10 dB 加性高斯白噪声的 Karl-Wilhelm-Strabe; snow 视频中的测试结果。因为基于稀疏分类器的特征算法对高斯白噪声的鲁棒性较高,所以本文算法和 AdaSR

算法在噪声情况下的跟踪效果较好。另外,AdaSR 算法因无需计算核函数空间的内积值,所以每帧耗时较本算法更低。

表 4 加入加性高斯白噪声的跟踪结果

Tab. 4 Tracking results for target added by white Gaussian noise

算法	误差 (50 次)	方差 (50 次)	每帧耗时	综合评价
本文算法	3.25	2.91	0.062	1.94
AdaSR	5.74	1.02	0.041	2.98
PF	11.21	2.18	0.052	5.84

综合上述实验结果可知,本文算法与 AdaSR 算法那比 PF 算法有更好的稳定性。本文提出的跟踪算法在多种复杂情况下能保证跟踪精度,虽然计算复杂度较其它两种算法有所增加,但总体能够保证运算的实时性。

7 结 论

本文针对经典稀疏分类目标跟踪算法在复杂情况下跟踪精度不高等问题,提出了在标准对冲框架下结合核函数稀疏分类 KRCD-SRC 算法与自适应字典更新方法的目标跟踪算法,该算法能够精确、稳定地跟踪目标。在使用 50 个粒子进行跟踪时,本文算法的处理帧率能够达到 14 frame/s。实验结果证明,本算法适用于复杂背景下的跟踪任务,在目标被遮挡或目标与背景具有相似特征模式情况下都能保持较高的跟踪精度。

参考文献:

- [1] JAWARD M, BULL D, CANAGARAJAH N. Sequential monte carlo methods for contour tracking of contaminant clouds[J]. *Signal Processing*, 2010, 90(1):249-260.
- [2] CHEN Q, SUN Q, HENG P A, XIA D S, et al. Two-stage object tracking method based on kernel and active contour[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2010, 20(4):605-609.
- [3] JING L, VADAKKEPAT P. Interacting mcmc particle filter for tracking maneuvering target[J]. *Dig-*

ital Signal Processing, 2010, 20(2):561-574.

- [4] HAN Z, JIAO J, ZHANG B, et al. Visual object tracking via sample-based adaptive sparse representation (AdaSR)[J]. *Pattern Recognition*, 2011, 44(9):2170-2183.
- [5] XUE M, LING H B. Robust visual tracking and vehicle classification via sparse representation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, 33(11):2259-2272.
- [6] WANG Q, CHEN F, XU X L, et al. Online discriminative object tracking with local sparse representation [C]. *IEEE Workshop on Applications of Computer Vision (WACV)*, 2012:425-432.

- [7] RIGAMONTI R, BROWN M A, LEPETIT V. Are sparse representations really relevant for image classification? [C]. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011: 1545-1552.
- [8] ZHANG L, YANG M, FENG X CH. Sparse representation or collaborative representation; Which helps face recognition? [J]. *IEEE International Conference on Computer Vision*, 2011:471-478.
- [9] HUANG K, AVIYENTE S. Sparse representation for signal classification [J]. *Advances in neural information processing*, 2007, 16(15): 609-617.
- [10] GAO SH G, IVOR W T, CHIA L T. Kernel sparse representation for image classification and face recognition [J]. *Lecture Notes in Computer Science*, 2010, 63(14): 1-14.
- [11] KANG C C, LIAO SH C, XIANG SH M, et al.. Kernel sparse representation with local patterns for face recognition [C]. *18th IEEE International Conference on Image Processing*, 2011:3009-3012.
- [12] 龚俊亮,何昕,魏仲慧,等.采用改进辅助粒子滤波的红外多目标跟踪[J]. *光学精密工程*, 2012, 20(2):413-421.
- GONG J L, HE X, WEI ZH H, et al.. Multiple infrared target tracking using improved auxiliary particle filter [J]. *Opt. Precision Eng.*, 2012, 20(2):413-421. (in Chinese)
- [13] 吴川,杨冬,郝志成.基于粒子滤波的彩色图像跟踪[J]. *光学精密工程*, 2009, 17(10):2542-2547.
- WU CH, YANG D, HAO ZH CH. Color image tracking algorithm based on particle filter [J]. *Opt. Precision Eng.*, 2009, 17(10):2542-2547. (in Chinese)
- [14] LI R, MAO L, ZHANG J R. Particle filter resampling based on optimized combinatorial algorithm [C]. *International Symposium on IT in Medicine and Education (ITME)*, 2011, 2:27-30.
- [15] CHAUDHURI K, FREUND Y, HSU D. A parameter-free hedging algorithm[C]. *Advances in Neural Information Processing Systems*, 2009: 297-305.
- [16] ETTINGER E, FREUND Y. Particle Filtering on the Audio Localization Manifold, *Arxiv preprint arXiv:1003.0659*, 2010.
- [17] DONOHO D L. For most large underdetermined systems of linear equations the minimal ℓ_1 -norm solution is also the sparsest solution. *Communications on Pure and Applied Mathematics* [J]. 2006, 59(6):797-829.
- [18] NESTEROV Y. Efficiency of coordinate descent methods on huge-scale optimization problems [J]. *SIAM Journal on Optimization*, 2012, 22(2):341-32.
- [19] PETER RICHTÁRIK, MARTIN TAKÁČ, Iteration complexity of randomized block-coordinate descent methods for minimizing a composite function [R]. *Arxiv preprint arXiv:1107.2848*, 2011.
- [20] 朱明清,王智灵,陈宗海.基于灰色预测模型和粒子滤波的视觉目标跟踪算法[J]. *控制与决策*, 2012, 27(1):53-57.
- ZHU M Q, WANG ZH L, CHEN Z H. Visual tracking algorithm based on grey prediction model and particle filter[J]. *Control and Decision*, 2012, 27(1):53-57. (in Chinese)

作者简介:



匡金骏(1984—),男,重庆人,重庆大学博士生,2010年赴美国加州大学圣迭戈分校联合培养,主要从事人工智能、光电目标追踪、机器视觉等方面的研究。E-mail:kuangjinjun@gmail.com



柴毅(1962—),男,安徽人,教授,博导,重庆大学自动化学院副院长,主要研究方向为信息处理,融合与控制,智能系统理论及其应用,工业工程控制理论与技术,计算机网络与系统控制。E-mail:chaiyi@cqu.edu.cn



熊庆宇(1963—),男,重庆人,教授,博导,重庆大学软件学院院长,中国仪器仪表学报评议专家,日本智能情报模糊学会会员,主要研究方向为智能感知、控制与决策等;E-mail:xiong03@cqu.edu.cn