

在线鲁棒判别式字典学习视觉跟踪

薛模根¹, 朱虹¹, 袁广林²

(1. 陆军军官学院偏振光成像探测技术安徽省重点实验室, 安徽合肥 230031; 2. 陆军军官学院十一系, 安徽合肥 230031)

摘要: 传统子空间跟踪较好解决了目标表观变化和遮挡问题, 但其仍存在对复杂背景下目标跟踪鲁棒性不足和模型漂移等问题. 针对这两个问题, 本文首先通过增大背景样本的重构误差和利用 L1 范数损失函数建立一种在线鲁棒判别式字典学习模型; 其次, 利用块坐标下降设计了该模型的在线学习算法用于视觉跟踪模板更新; 最后, 以粒子滤波为框架, 结合提出的模板更新方法实现了鲁棒的视觉跟踪. 实验结果表明: 与 IVT (Incremental Visual Tracking)、LIAPG (L1-tracker using Accelerated Proximal Gradient)、ONNDL (Online Non-Negative Dictionary Learning) 和 PCOM (Probability Continuous Outlier Model) 等典型跟踪方法相比, 本文方法具有较强的鲁棒性和较高的跟踪精度.

关键词: 视觉跟踪; 模板更新; 字典学习; 粒子滤波

中图分类号: TP391.4

文献标识码: A

文章编号: 0372-2112 (2016)04-0838-08

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.3969/j.issn.0372-2112.2016.04.012

Online Robust Discrimination Dictionary Learning for Visual Tracking

XUE Mo-gen¹, ZHU Hong¹, YUAN Guang-lin²

(1. Anhui Province Key Laboratory of Polarization Imaging Detection Technology, Army Officer Academy of PLA, Hefei, Anhui 230031, China;

(2. Eleventh Department, Army Officer Academy of PLA, Hefei, Anhui 230031, China)

Abstract: The traditional subspaces based visual trackers well solved appearance changes and occlusions. However, they were weakly robust for complex background and prone to model drifting. To deal with these two problems, this paper enlarges reconstruction errors of the background samples and uses L1-norm loss function to establish an online robust discrimination dictionary learning model. Then an online robust discrimination dictionary learning algorithm for template updating in visual tracking is designed via the block coordinate descent (BCD). Finally, robust visual tracking is achieved with the proposed template updating method in particle filter framework. The experimental results show that the proposed method has better performance in robustness and accuracy than the state-of-the-art trackers such as IVT (Incremental Visual Tracking), LIAPG (L1-tracker using Accelerated Proximal Gradient), ONNDL (Online Non-Negative Dictionary Learning) and PCOM (Probability Continuous Outlier Model).

Key words: visual tracking; template updating; dictionary learning; particle filter

1 引言

视觉跟踪是通过视频图像序列不断估计目标状态的过程, 它在智能监控、人机交互、机器人导航和运动分析等方面都具有重要的应用价值. 目前, 视觉跟踪面临着目标表观变化、目标遮挡和复杂背景等难题, 其仍是计算机视觉领域研究的热点. 为解决这些难题, 已经提出了一些方法^[1], 其中, 基于子空间的跟踪方法较好解决了目标表观变化和遮挡问题, 特别是 L1 跟踪的

提出使子空间跟踪再次受到了广泛关注.

受到主成分分析 (Principal Component Analysis, PCA) 在人脸识别中应用的启发, 1996 年, Black 等人^[2]首次将基于 PCA 的子空间模型应用到视觉跟踪中. 该方法以子空间常量为假设, 不能适应目标表观的变化. 为了适应目标的变化, 子空间更新变得尤为重要, 因此 Ross 等人^[3]提出一种基于增量子空间学习 (Incremental Subspace Learning, ISL) 的视觉跟踪. 该方法利用 ISL 更新子空间, 适应了目标表观的缓慢变化, 但是易于发生

模型漂移,对目标遮挡的鲁棒性较差.近几年,受到稀疏表示在人脸识别中应用的启发,一些学者提出了基于稀疏表示的子空间跟踪^[4-7],又称为“L1跟踪”.L1跟踪较好地解决了目标遮挡问题,但它对模板的依赖性更强,一旦模板中引入错误信息容易导致跟踪失败,这就对模板更新提出了更高的要求.文献[4,5]首先检测目标是否受到遮挡,并利用未遮挡的跟踪结果直接替换模板,一定程度上缓解了模型漂移.文献[6]提出一种基于在线字典学习的目标跟踪方法,利用字典学习在线更新模板,增强了对目标变化的适应性和鲁棒性.文献[7]将在线鲁棒字典学习和非负矩阵分解相结合用于模板更新,有效抑制了模型漂移.在L1跟踪的启发下,Wang等人^[8]在文献[3]基础上利用PCOM描述连续离群点用于遮挡检测,进一步提高了对遮挡目标跟踪的鲁棒性.

由上述分析可以发现,现有子空间跟踪较好解决了目标表观变化和目标遮挡问题,但仍存在一些不足:对于复杂背景下的目标跟踪,由于产生式模型中缺少判别信息,模板很难对相似的目标和背景进行准确判别,从而导致跟踪失败.针对此问题,本文首先提出一种在线鲁棒判别式字典学习模型,该模型引入背景信息并增大模板对背景样本的重构误差,提高了模板的判别能力;同时,为了克服模型漂移,采用L1范数作为损失函数,降低了离群数据(如错误信息或遮挡信息)对模板的影响.然后,求解该模型提出了在线鲁棒判别式字典学习算法用于视觉跟踪模板更新.

2 在线鲁棒判别式字典学习

由文献[6,7]可知,视觉跟踪模板更新可以看作是在线的字典学习问题.与一般在线字典学习不同,本文结合视觉跟踪中对模板的判别性和鲁棒性要求,提出了在线鲁棒判别式字典学习模型,并设计了其求解算法用于视觉跟踪模板更新.

2.1 在线鲁棒判别式字典学习模型

为了使模板包含判别式信息同时抑制模型漂移,本文提出如式(1)所示的在线鲁棒判别式字典学习模型:

$$\min_{(D, X)} \frac{1}{nN} \sum_{i=1}^n (\|A_i^+ - DX_i^+\|_1 - \lambda_1 \|A_i^- - DX_i^-\|_1 + \lambda_2 \|X_i\|_F^2) \quad \text{s. t. } \|d\|_2 = 1, \forall d \quad (1)$$

其中,训练样本集 $\{A_1, \dots, A_n\}$ 由 n 个样本子集组成,每个样本子集 $A_i (i=1, \dots, n)$ 包含目标样本 A_i^+ 和背景样本 A_i^- 共 N 个训练样本; D 为目标模板字典; $X_i = [X_i^+, X_i^-]$, X_i^+, X_i^- 分别为 A_i^+, A_i^- 在 D 上的编码系数, $i=1, \dots, n$; d 为 D 上任意原子; λ_1, λ_2 为调节常数.

本文提出的在线鲁棒判别式字典学习模型式(1)包含损失函数和正则化项两个部分.在字典学习中,普遍采用的损失函数有L1损失函数和L2损失函数.与L2损失函数相比,L1损失函数具有对离群数据鲁棒的优点,即利用L1损失函数学到的字典原子受离群数据的影响较小.这一结论已在人脸识别^[9]和背景估计^[10]等计算机视觉中得到证实.其原因在于,离群数据的重构误差满足拉普拉斯分布,这正好与L1损失相符^[11].因此,为了降低离群数据对目标模板的影响,从而抑制模型漂移,模型式(1)采用L1范数作为损失函数.为了适应目标的变化,模板字典 D 对目标样本应具有较好地重构能力,目标样本的重构误差越小越好,即 $\min \|A_i^+ - DX_i^+\|_1$.为了增强模板字典 D 对目标和背景的判别能力,背景样本的重构误差越大越好,即 $\max \|A_i^- - DX_i^-\|_1$.综合以上三个方面,损失函数应为 $\min(\|A_i^+ - DX_i^+\|_1 - \lambda_1 \|A_i^- - DX_i^-\|_1)$,其中 λ_1 用于平衡两个重构误差. $\lambda_2 \|X_i\|_F^2$ 为正则化项,使目标函数的解更稳定.

2.2 在线鲁棒判别式字典学习算法

为了适应目标表观的变化,在视觉跟踪中需要利用跟踪结果更新目标模板.显而易见在线字典学习的学习方式比较符合视觉跟踪的要求.启发于文献[12],求解如式(1)所示模型的在线鲁棒判别式字典学习算法分为三个阶段:采集在线数据、求解编码系数和在线字典更新.假设每隔 h 帧图像序列进行一次模板更新, t 时刻第 T 次更新时, $n=T$,采集在线数据阶段,取 $t-h+1, \dots, t$ 时刻的跟踪结果作为目标样本 A_t^+ ;取以 t 时刻跟踪结果中心位置 l 为圆心的环形区域 $\{l' | \gamma < \|l' - l\| < \delta\}$ 内随机值作为背景样本 A_t^- .求解编码系数阶段,已知 $t-h$ 时刻模板字典 D_{t-h} 以及训练样本 A_t^+, A_t^- ,采用迭代重加权最小平方法(Iterative Reweighted Least Squares, IRLS)^[13]分别求解式(2)、式(3)得到 X_t^+, X_t^- :

$$X_t^+ = \arg \min_{X_t^+} \|A_t^+ - D_{t-h} X_t^+\|_1 + \lambda_2 \|X_t^+\|_F^2 \quad (2)$$

$$X_t^- = \arg \max_{X_t^-} \lambda_1 \|A_t^- - D_{t-h} X_t^-\|_1 - \lambda_2 \|X_t^-\|_F^2 \quad (3)$$

在线字典更新阶段,已知编码系数 X_t^+, X_t^- 求解模板字典 D_t ,如式(4)所示:

$$D_t = \arg \min_D \frac{1}{TN} \sum_{i=1}^T (\|A_i^+ - DX_i^+\|_1 - \lambda_1 \|A_i^- - DX_i^-\|_1) \quad (4)$$

$$= \arg \min_D \frac{1}{TN} (\|A^+ - DX^+\|_1 - \lambda_1 \|A^- - DX^-\|_1) \quad (5)$$

其中, $A^+ = [A_1^+, \dots, A_T^+]$, $A^- = [A_1^-, \dots, A_T^-]$, $X^+ =$

$[X_1^+, \dots, X_r^+], X^- = [X_1^-, \dots, X_r^-]$. 鉴于对跟踪效率的考虑, 采用块坐标下降法 BCD^[14] 逐行迭代求解式(4). 将式(5)转化为行优化函数, 如式(6)所示:

$$D_t(j, :) = \arg \min_d \frac{1}{TN} \left(\sum_{i=1}^{N^+} |(A^+)_{ji} - d(X^+)_{i}| \right) - \lambda_1 \sum_{i=1}^{N^-} |(A^-)_{ji} - d(X^-)_{i}| \quad (6)$$

其中, d 为 D 的第 j 行向量 ($j=1, \dots, k$), $(A^+)_{ji}, (A^-)_{ji}$ 分别为 A^+, A^- 的第 j 行第 i 列元素, $(X^+)_{i}, (X^-)_{i}$ 分别为 X^+, X^- 的第 i 列向量, N^+, N^- 分别为目标样本数和背景样本数. 求解式(6)可以通过式(7)、式(8)不断迭代逼近其解直至收敛:

$$D_t(j, :) = \arg \min_d \frac{1}{TN} \left(\sum_{i=1}^{N^+} w_{ji}^+ ((A^+)_{ji} - d(X^+)_{i})^2 \right) - \lambda_1 \sum_{i=1}^{N^-} w_{ji}^- ((A^-)_{ji} - d(X^-)_{i})^2 \quad (7)$$

$$\begin{cases} w_{ji}^+ = \frac{1}{\sqrt{((A^+)_{ji} - D_t(j, :)(X^+)_{i})^2 + \delta^+}} \\ w_{ji}^- = \frac{1}{\sqrt{((A^-)_{ji} - D_t(j, :)(X^-)_{i})^2 + \delta^-}} \end{cases} \quad (8)$$

其中, δ^+, δ^- 为很小的正数. 显然, 每次迭代求解式(7)即是最优化一个二次函数, 易于解决. 通过上述研究分析, 本文提出的基于在线鲁棒判别式字典学习的模板更新方法总结如算法 1 所示.

算法 1 模板更新方法

- 1: 输入: $t-h$ 时刻, $n = T-1$: 模板字典 D_{t-h} , 数据记录 $B_{T-1}^j, C_{T-1}^j, j=1, \dots, k$
- 2: t 时刻, $n = T$:
- 3: 采集在线数据: 目标样本 A_t^+ 和背景样本 A_t^-
- 4: 求解编码系数: 采用 IRLS 分别求解式(2)、式(3)得到 X_t^+, X_t^-
- 5: 在线字典更新:
- 6: 初始化: $\begin{cases} w_{ji}^+ = 1, (j=1, \dots, k; i=1, \dots, N^+) \\ w_{ji}^- = 1, (j=1, \dots, k; i=1, \dots, N^-) \end{cases}$
- 7: 重复
- 8: for $j=1$ to k do
- 9: $\begin{cases} B_t^j = B_{T-1}^j + \sum_{i=1}^{N^+} w_{ji}^+ (A_t^+)_{ji} (X_t^+)_{i}^T - \sum_{i=1}^{N^-} \lambda_1 w_{ji}^- (A_t^-)_{ji} (X_t^-)_{i}^T \\ C_t^j = C_{T-1}^j + \sum_{i=1}^{N^+} w_{ji}^+ (X_t^+)_{i} (X_t^+)_{i}^T - \sum_{i=1}^{N^-} \lambda_1 w_{ji}^- (X_t^-)_{i} (X_t^-)_{i}^T \end{cases}$
- 10: 求解线性系统: $B_t^j = D_t(j, :) C_t^j$
- 11: 计算权值: $\begin{cases} w_{ji}^+ = \frac{1}{\sqrt{((A_t^+)_{ji} - D_t(j, :)(X_t^+)_{i})^2 + \delta^+}}, (i=1, \dots, N^+) \\ w_{ji}^- = \frac{1}{\sqrt{((A_t^-)_{ji} - D_t(j, :)(X_t^-)_{i})^2 + \delta^-}}, (i=1, \dots, N^-) \end{cases}$
- 12: end for

13: 归一化处理

14: 直到收敛

15: 输出: $D_t, B_t^j, C_t^j, j=1, \dots, k$

3 目标跟踪

本文跟踪方法是以粒子滤波为框架结合提出的模板更新方法而建立的. 利用粒子滤波跟踪目标, 包括预测和更新两个步骤. 已知 1 到 $t-1$ 时刻所有可用图像观测 $y_{1:t-1} = (y_1, \dots, y_{t-1})$, 则预测过程为:

$$p(x_t | y_{1:t-1}) = \int p(x_t | x_{t-1}) p(x_{t-1} | y_{1:t-1}) dx_{t-1} \quad (9)$$

其中, x_t, x_{t-1} 分别为 t 和 $t-1$ 时刻的目标状态, $p(x_t | x_{t-1})$ 为状态转移模型. 本文采用高斯分布建立状态转移模型, 如式(10)所示:

$$p(x_t | x_{t-1}) = N(x_t; x_{t-1}, \Psi) \quad (10)$$

其中, $x_t = (x_t, y_t, w_t, h_t, \theta_t)$, $x_t, y_t, w_t, h_t, \theta_t$ 分别表示 t 时刻目标区域中心点位置坐标、目标宽度、高度及倾斜角. Ψ 为对角矩阵, 其对角元素表示相应状态的方差. t 时刻, 当图像观测 y_t 可用时, 则进行更新过程:

$$p(x_t | y_{1:t}) \propto p(y_t | x_t) p(x_t | y_{1:t-1}) \quad (11)$$

其中, $p(y_t | x_t)$ 为观测似然模型. 建立观测似然模型分为两个步骤, 首先, 对于任意粒子的图像观测 y_t^i , 求解其 L1 正则化编码系数, 模型如式(12)所示:

$$\min_{e_i} \| B_t c_i^j - y_t^i \|_2 + \lambda \| c_i^j \|_1 \quad (12)$$

其中, $B_t = [D_t, I]$, D_t 为 t 时刻目标模板, I 是一个单位阵, 称为小模板; 相应地, $(c_i^j)^T = [a_i^j, e_i^j]$, a_i^j 为图像观测 y_t^i 在目标模板 D_t 上的稀疏编码系数, e_i^j 为图像观测 y_t^i 在小模板 I 上的稀疏编码系数, 称为小模板系数; λ 为正则化系数. 然后, 已知编码系数定义如式(13)所示观测似然模型:

$$p(y_t^i | x_t^i) = \frac{1}{\Gamma} \exp(-\alpha \| D_t a_i^j - y_t^i \|_2^2) \quad (13)$$

其中, Γ 为归一化常量, α 为高斯核尺度参数. 需要特别说明的是, 本文是在 L1 跟踪基础上的研究工作, L1 跟踪最大的优势是在目标遮挡下能够鲁棒跟踪, 其原因在于目标遮挡使观测噪声符合拉普拉斯分布. 据此, L1 跟踪利用小模板 I 表示噪声以捕捉遮挡, 如式(12)模型所示. 综上所述, 已知所有图像观测 $y_{1:t}$, 利用最大后验概率准则估计跟踪目标的最优状态:

$$\begin{aligned} \hat{x}_t &= \arg \max_x p(x_t^i | y_{1:t}) \\ &= \arg \max_x p(y_t^i | x_t^i) p(x_t^i | y_{1:t-1}), i=1, 2, \dots, N' \end{aligned} \quad (14)$$

其中, x_t^i 为第 i 个粒子在 t 时刻的状态, N' 为粒子数. 以粒子滤波为框架, 结合状态转移模型、观测似然模型和

模板更新方法建立目标跟踪方法,如算法 2 所示.

算法 2 目标跟踪方法

- 1: 初始化: $t=1$ 时刻,手工标定目标 x_1 , 初始目标模板 D_1 , 初始粒子权值 $1/N'$, $h=1$
- 2: for $t=2$ to T do
- 3: 采样粒子: 利用状态转移模型式(10)预测粒子状态 x_t^i , 采样图像观测 y_t^i , $i=1, 2, \dots, N'$
- 4: 稀疏编码: 对于所有 y_t^i , 求解式(12)得到编码系数 c_t^i , $i=1, 2, \dots, N'$
- 5: 权值更新: 首先利用式(13)计算观测似然 $p(y_t^i | x_t^i)$, 再用 $w_t^i = w_{t-1}^i p(y_t^i | x_t^i)$ 更新权值, $i=1, 2, \dots, N'$
- 6: 权值归一化: $w_t^i = w_t^i / \sum_{i=1}^{N'} w_t^i$, $i=1, 2, \dots, N'$
- 7: 目标状态估计: $\hat{x}_t = x_t^{\hat{i}}$, 其中 $\hat{i} = \max_i(w_t^i)$
- 8: 粒子重采样
- 9: $h = h + 1$
- 10: 如果 h 达到规定阈值, 则利用算法 1 更新目标模板, 并且重置 $h=0$
- 11: 输出: 跟踪结果 \hat{x}_t
- 12: end for

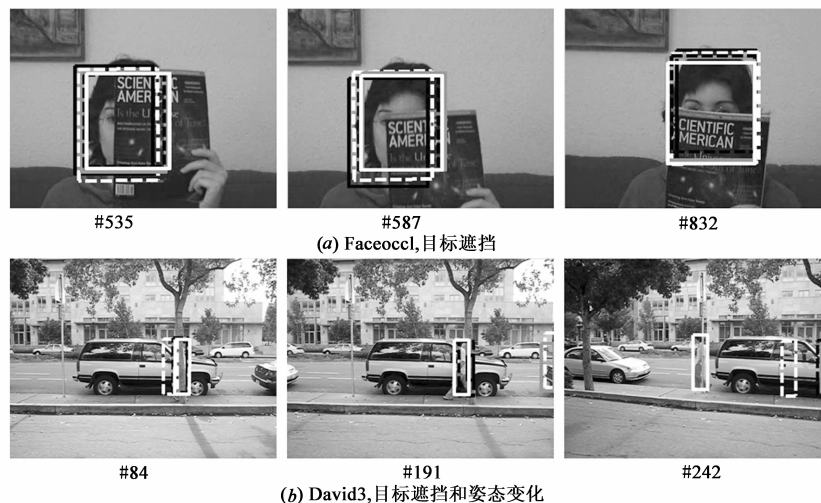
4 实验结果与分析

本文提出的跟踪方法是以 Matlab R2010a 为开发工具实现的, 并在 Intel (R) Core (TM) 3. 19GHz 4CPUs, 3. 47GB 内存的台式电脑上调试通过. 采用 Faceoccl、David3、Walking2、Singer1、Dog1、Suv、Cardark 等序列^[15]对提出的跟踪方法进行了实验验证, 并与 IVT^[3]、LIAPG^[5]、ONNDL^[7]、PCOM^[8]等 4 种跟踪方法进行了比较. 为了保证比较的合理性和公平性, 实验中, 5 种跟

踪方法的粒子数均为 600, 模板(特征基)的大小均为 32×32 , 个数均为 16, 状态转移模型参数均为 $(4, 4, 0.01, 0.01, 0.002)$. 关于本文跟踪方法的参数设置说明如下: 在线鲁棒判别式字典学习算法中, 参数 λ_1 用于平衡目标和背景重构误差, 参数 λ_2 为正则化系数, 用于平衡损失函数与正则化项, 为了保证目标表示的精准度, 本文设置 λ_1, λ_2 为较小的正数, 通过选取 $0 \sim 1$ 之间的正数进行实验发现当 $\lambda_1 = 0.25, \lambda_2 = 0.01$ 时跟踪性能最好; 模板更新步长 h 取决于目标表观变化频率, 本文设置 $h=5$, 即每隔 5 帧图像序列进行一次模板更新; 采集在线数据时, 取前 5 帧跟踪结果为目标样本 ($N^+ = 5$), 取当前帧环形区域内随机值为背景样本 ($N^- = 200$); 参数 γ 和 δ 为环形区域的两个半径, 其值的选取应使背景样本包含大量背景信息和少量目标信息, 本文采用与文献[16]相同的方法设置参数 γ 和 δ , 即 $\gamma = \max(W/2, H/2), \delta = 2\gamma$, 其中 W 和 H 分别为目标的宽和高.

4.1 定性分析

图 1 给出了 IVT、LIAPG、ONNDL、PCOM 及本文跟踪方法对 7 组图像序列的跟踪结果和定性比较. Faceoccl、David3、Walking2 和 Suv 序列存在大面积目标遮挡, 由图 1(a)~(c)、1(f)可知, 本文跟踪方法成功抑制了模型漂移, 明显优于其他跟踪方法. David3、Singer1 和 Dog1 序列存在大量目标表观变化, 由图 1(b)、1(d)、1(e)可知, 本文跟踪方法优于其他 4 种方法, 适应性好. Dog1、Suv 和 Cardark 序列受到复杂或低对比度背景干扰, 由图 1(e)~(g)可知, 本文跟踪方法的跟踪结果最优, 具有强判别力.



(b) David3, 目标遮挡和姿态变化

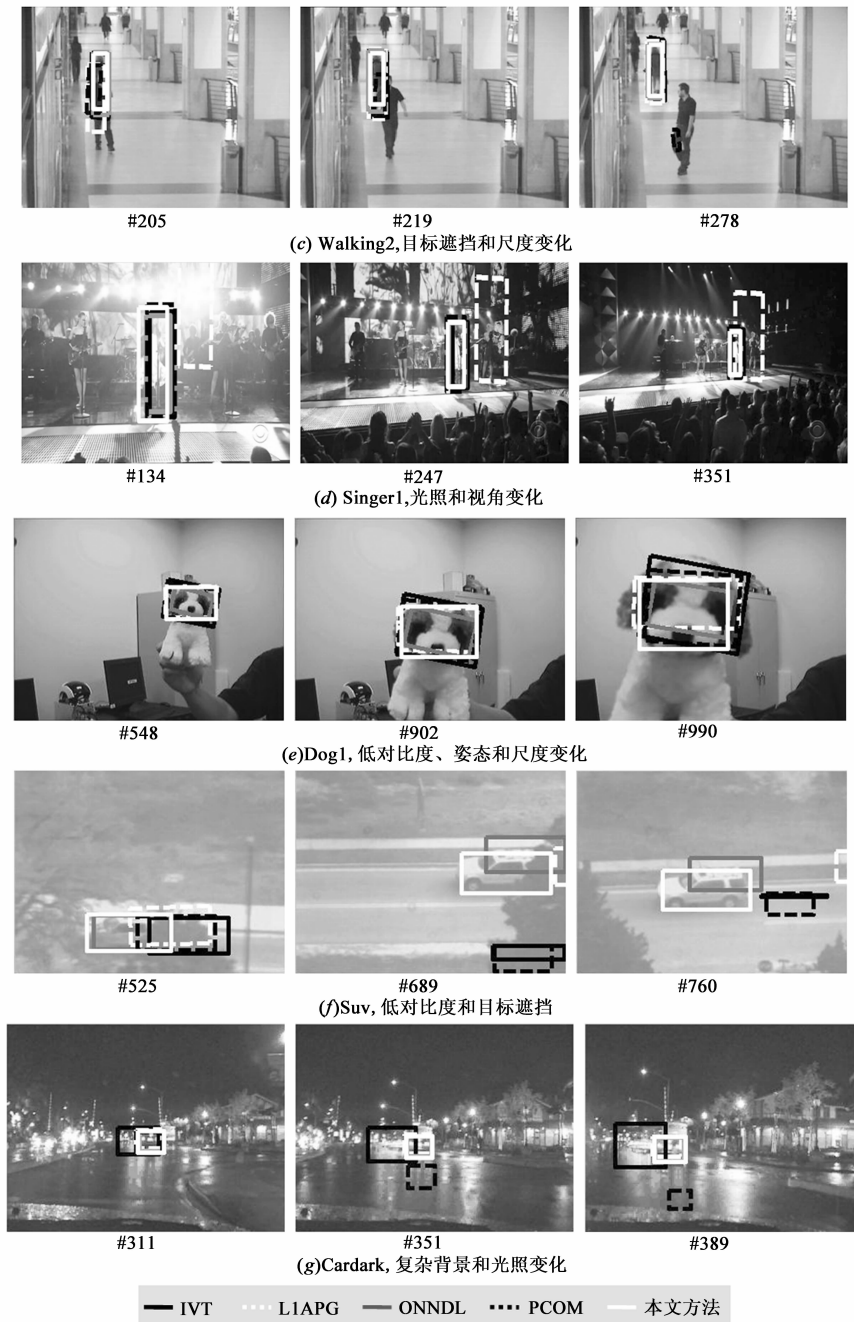


图1 5种方法的跟踪结果比较

4.2 定量分析

借鉴于文献[15], 本文采用成功率和精度两个指标进行定量实验分析. 给定一帧图像, 已知某种跟踪方法的跟踪窗区域 r_i 和实际跟踪窗区域 r_a , 定义重叠率为 $S = |r_i \cap r_a| / |r_i \cup r_a|$, 其中, $|\cdot|$ 表示区域内像素个数; 定义中心位置误差为 r_i 与 r_a 中心点之间的欧式距离(像素个数). 据此, 关于两个实验指标的定义及参数

设置说明如下: 成功率为重叠率大于给定重叠阈值的图像帧数比率, 实验中设定重叠阈值为 0 到 1. 精度为中心位置误差小于给定中心位置误差阈值的图像帧数比率, 实验中设定中心位置误差阈值为 0 到 50 个像素. 图 2 给出了 5 种跟踪方法对 7 组图像序列的成功率和精度曲线比较. 由图 2 可知, 本文方法在跟踪成功率和精度上均优于其他方法.

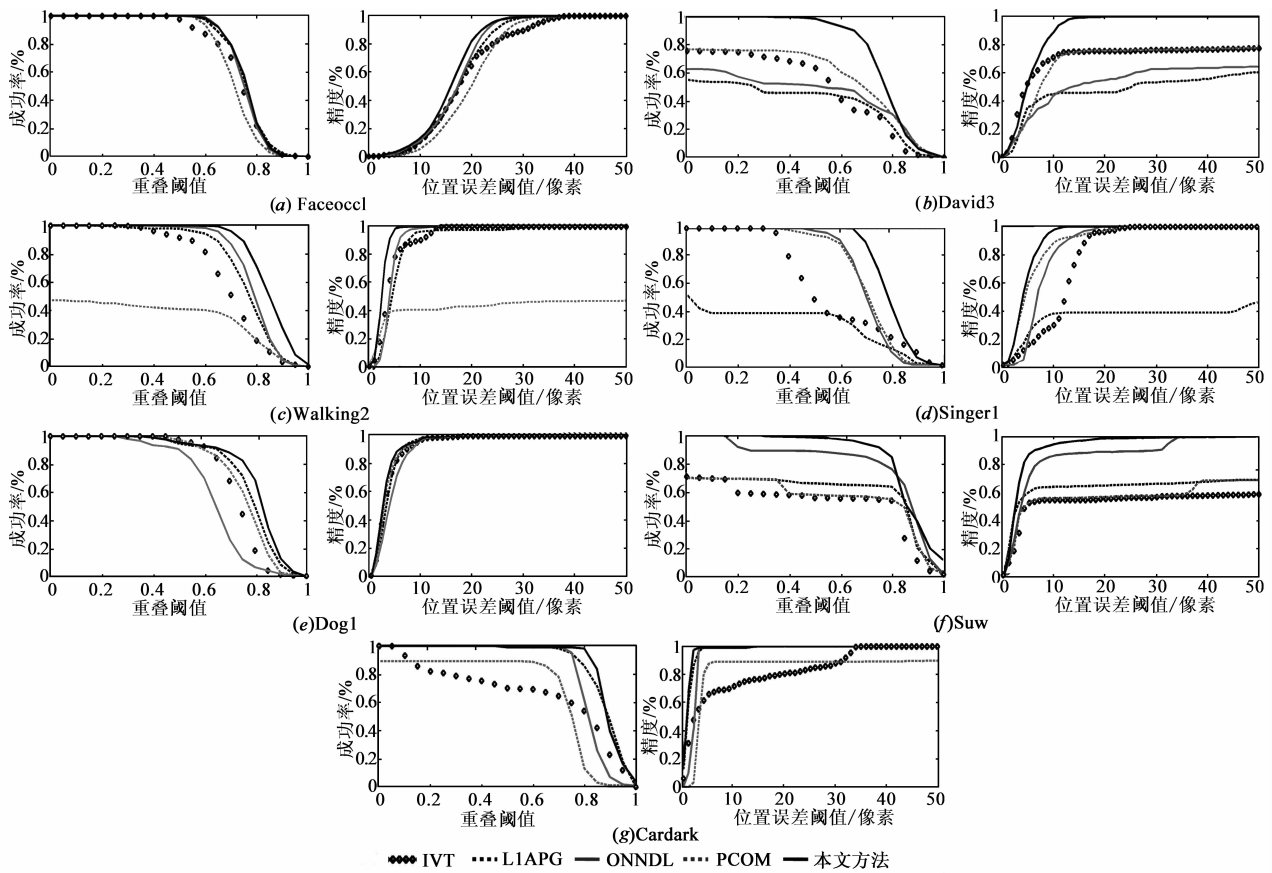


图2 5种方法成功率曲线和精度曲线比较

4.3 计算时间分析

实验中作比较的 5 种跟踪方法 (IVT、LIAPG、ONNDL、PCOM 及本文方法) 均是以粒子滤波为框架的子空间跟踪. 假设 $U \in R^{k \times q}$ 为 IVT 和 PCOM 的特征基, $D \in R^{k \times q}$ 为 LIAPG、ONNDL 和本文方法的模板, N' 为粒子数. 在相同软硬件环境下, $k = 32 \times 32, q = 16, N' = 600$ 时, 5 种跟踪方法进行一帧跟踪的平均计算时间如表 1 所示. 可以看出, 本文跟踪方法的速度比 IVT、PCOM 慢, 但比 LIAPG、ONNDL 快.

表 1 5 种方法的计算时间比较

跟踪方法	IVT	LIAPG	ONNDL	PCOM	Our
计算时间(s)	0.038	0.678	2.646	0.072	0.192

4.4 鲁棒性分析

对比 LIAPG、ONNDL 与本文方法相似的跟踪方法, 由上述实验可知, 本文方法具有较强的鲁棒性与较高的跟踪精度, 其根本原因在于本文跟踪方法与它们使用的模板更新方法不同. 本文提出了一种在线鲁棒判别式字典学习算法进行模板更新. 一方面, 该算法增强模板对目标和背景的辨识度, 实现了判别式字典学习, 提高了跟踪精度; 另一方面, 采用 L1 范数损失函数和在线学习方法, 在适应目标变化的同时克服了模型漂移, 进一步提高了跟踪鲁棒性. 图 3 分别给出了 Faceoccl 序列中 LIAPG、ONNDL 和本文方法在第 350 帧时模板更新结果. 可以看出, 本文模板不仅自适应目标变化, 还成功排除了遮挡信息, 完好保留了目标信息.

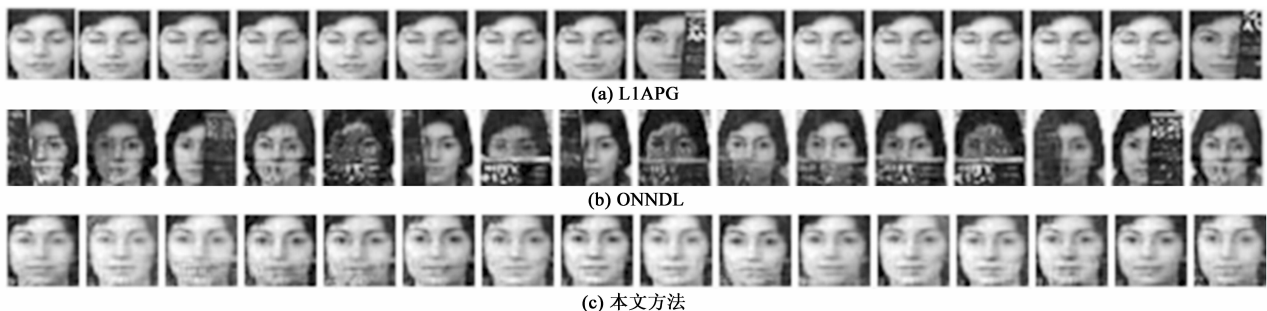


图3 第350帧时LIAPG、ONNDL和本文方法的模板更新结果

4.5 讨论

上述实验表明,本文跟踪方法具有较强的鲁棒性和较高的跟踪精度,但它还存在两个问题.第一,与现有基于字典学习的跟踪方法^[6,7]相同,利用在线字典学习得到的目标模板字典不完整和不准确,尤其是在目标外观突然变化时,比如快速 out-plane 旋转和突然的光照变化等情况,会导致跟踪失败.图 4(a)给出了本文方法存在跟踪失败的情况. Girl 序列的跟踪目标为人头,它存在快速的 out-plane 旋转,发生了跟踪失败.第

二,本文跟踪方法是利用粒子滤波实现的.众所周知,粒子滤波较好解决了目标的非线性运动问题,但是在目标快速运动时,基于粒子滤波的跟踪方法需要大规模的粒子才能稳定跟踪目标,这会显著降低目标跟踪的速度.图 4(b)给出了当目标快速运动时,本文方法的跟踪结果.虽然能够稳定跟踪目标,但是其跟踪速度较慢,平均每帧的跟踪时间为 0.556s.上述两个方面是进一步研究要解决的问题.

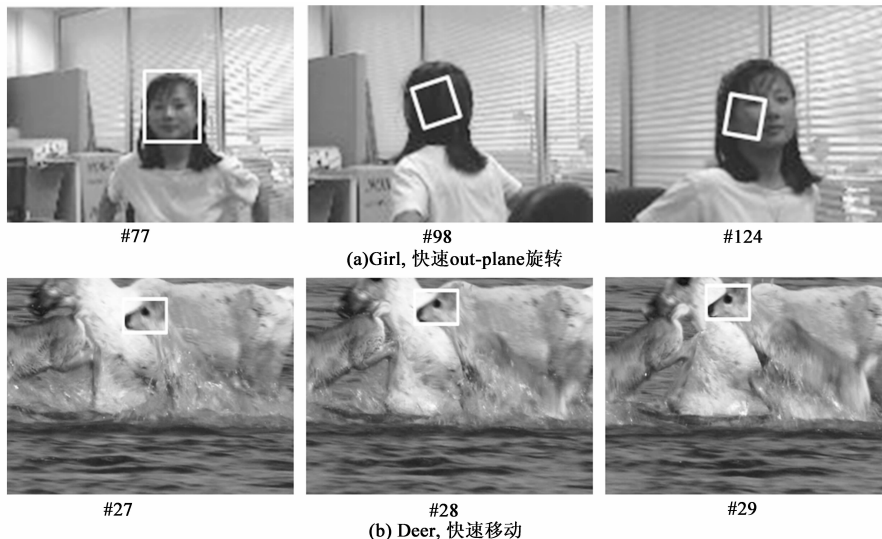


图4 两个进一步研究的问题

5 结论

针对现有子空间跟踪对复杂背景下目标跟踪判别力不强和模型漂移的问题,本文首先提出了一种在线鲁棒判别式字典学习模型.一方面,该模型通过增大模板重构背景样本的误差提高了模板对目标和背景的判别能力;另一方面,使用 L1 范数作为损失函数降低了模板对离群数据的敏感度.然后,求解该模型设计了在线学习算法用于视觉跟踪模板更新.在上述两个方面的基础上,以粒子滤波为框架完成了在线鲁棒判别式字典学习视觉跟踪.利用多个具有挑战性的图像序列对提出的跟踪方法进行了实验验证并与现有跟踪方法进行了比较,实验结果表明:与现有跟踪方法相比,本文方法更能鲁棒跟踪目标.

参考文献

[1] Li Xi, Hu Wei-ming, Shen Chun-hua, et al. A survey of appearance models in visual object tracking[J]. ACM Transactions on Intelligent Systems and Technology, 2013, 4(4):58(1) - 58(48).

[2] Black M J, Jepson A D. Eigentracking: Robust matching and tracking of articulated objects using a view-based rep-

resentation[A]. European Conference on Computer Vision [C]. London: Springer-Verlag Berlin, 1996. 329 - 342.

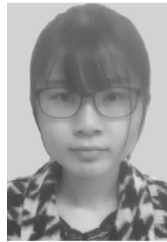
- [3] Ross D, Lim J, Lin R S, et al. Incremental learning for robust visual tracking[J]. International Journal of Computer Vision, 2008, 77(1 - 3):125 - 141.
- [4] Mei Xue, Ling Hai-bin, Wu Yi, et al. Minimum error bounded efficient L1 tracker with occlusion detection[A]. IEEE Conference on Computer Vision and Pattern Recognition [C]. Colorado: Springer-Verlag Berlin, 2011. 1257 - 1264.
- [5] Bao Cheng-long, Wu Yi, Ling Hai-bin, et al. Real time robust L1 tracker using accelerated proximal gradient approach[A]. IEEE Conference on Computer Vision and Pattern Recognition [C]. Providence: IEEE Computer Society Press, 2012. 1830 - 1837.
- [6] Xing Jun-liang, Gao Jin, Li Bing, et al. Robust object tracking with online multi-lifespan dictionary learning [A]. IEEE International Conference on Computer Vision [C]. Sydney: IEEE Computer Society Press, 2013. 665 - 672.
- [7] Wang Nai-yan, Wang Jing-dong, Yeung D. Online robust non-negative dictionary learning for visual tracking [A]. IEEE International Conference on Computer Vision [C]. Sydney: IEEE Computer Society Press, 2013. 657 - 664.

- [8] Wang Dong, Lu Hu-chuan. Visual tracking via probability continuous outlier model [A]. IEEE Conference on Computer Vision and Pattern Recognition [C]. Columbus: IEEE Computer Society Press, 2014. 3478 – 3485.
- [9] Wagner A, Wright J, Ganesh A, et al. Towards a practical face recognition system: Robust alignment and illumination by sparse representation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34 (2): 372 – 386.
- [10] Lu Ce-wu, Shi Jian-ping, Jia Jia-ya. Online robust dictionary learning [A]. IEEE Conference on Computer Vision and Pattern Recognition [C]. Portland: IEEE Computer Society Press, 2013. 415 – 422.
- [11] Nojun K. Principal component analysis based on L1-norm maximization [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 30(9): 1672 – 1680.
- [12] Mairal J, Bach F, Ponce J, et al. Online dictionary learning for sparse coding [A]. The 26th International Conference on Machine Learning [C]. Montreal: IEEE Computer Society Press, 2009. 539 – 547.
- [13] Bissantz N, Dmbgen L, Munk A, et al. Convergence analysis of generalized iteratively reweighted least squares algorithms on convex function spaces [J]. SIAM Journal on Optimization, 2009, 19: 1828 – 1845.
- [14] Richtarik P, Takac M. Iteration complexity of randomized block-coordinate decent methods for minimizing a composite function [J]. Mathematical Programming, 2014, 144 (1 – 2): 1 – 38.
- [15] Wu Yi, Lim J, Yang M. Online object tracking: a benchmark [A]. IEEE Conference on Computer Vision and Pattern Recognition [C]. Portland: IEEE Computer Society Press, 2013. 2411 – 2418.
- [16] Babenko B, Yang M H, Belongie S. Robust object tracking with online multiple instance learning [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(8): 1619 – 1632.

作者简介



薛模根 男, 1964 年 10 月出生于安徽合肥. 博士, 教授, 合肥工业大学博士生导师. 主要研究方向为图像处理、计算机视觉、光电防御等.
E-mail: xuemogen@126.com



朱虹 女, 1987 年 10 月出生于新疆乌鲁木齐. 现为陆军军官学院硕士研究生. 主要研究方向为图像处理、计算机视觉等.
E-mail: 729039126@qq.com