

文章编号 1004-924X(2018)04-0989-09

基于稀疏正则约束的子空间视觉跟踪

陈典兵^{1,2}, 朱 明^{1*}, 王慧利^{1,2}, 杨 航¹

(1. 中国科学院 长春光学精密机械与物理研究所, 长春 130033; 2. 中国科学院大学, 北京 100039)

摘要: 针对子空间表示跟踪算法处理遮挡问题的能力不足, 以及稀疏表示跟踪算法无法满足跟踪实时性要求等问题, 本文提出一种稀疏正则约束的子空间视觉跟踪算法。该算法结合了子空间表示与稀疏表示的优势, 提升了对于遮挡问题的处理能力, 并且降低了算法的计算复杂度。首先, 该算法利用 PCA 子空间基向量集、子空间均值以及表示残差对目标进行表示, 同时算法采用 L_2 范数作为表示系数以及表示残差的稀疏约束函数。其次, 算法采用了一种分步循环迭代的方法求解表示模型的系数与残差。然后, 为了保证子空间基向量与空间均值能够持续准确的描述目标在跟踪过程中出现的变化, 算法根据经过开运算处理后的表示残差中非零元素的不同比率构建不同的更新模板, 并结合增量主成分分析方法在线学习新的基向量与均值。最后, 在实验部分, 本文将提出算法在 10 个实验序列上的跟踪结果与 8 个现今主流跟踪算法进行对比, 同时从定性及定量两个方面对实验结果进行分析。本文算法在全部 10 个实验序列上的平均中心误差为 5.3 pixel, 平均覆盖率为 77%, 相比于其他算法, 本文算法取得了较高的跟踪精度。本文算法具有更好的鲁棒性, 并且满足更多场景下的跟踪需求。

关键词: 稀疏表示; 视觉跟踪; 子空间表示; 粒子滤波

中图分类号: TP391.41 **文献标识码:** A **doi:** 10.3788/OPE.20182604.0989

Visual tracking algorithm based on the sparse-regularized subspace

CHEN Dian-bing^{1,2}, ZHU Ming^{1*}, WANG Hui-li^{1,2}, YANG Hang¹

(1. *Changchun Institute of Optics, Fine Mechanics and Physics,
Chinese Academy of Sciences, Changchun 130033, China;*

2. *University of Chinese Academy of Sciences, Beijing 100039, China)*

** Corresponding author, E-mail: zhu_mingca@163.com*

Abstract: Aiming at the problem that the subspace representation tracking algorithm cannot deal with the occlusion problem effectively and the sparse representation tracking algorithm cannot meet the real-time requirements of the tracking, this paper proposed a sparse-regularized subspace visual tracking algorithm. The algorithm combined the advantages of subspace representation and sparse representation, improved the processing ability of the occlusion problem, and reduced the computational complexity. Firstly, the algorithm adopted the PCA subspace basis, the subspace mean and the representation residual to represent the target, and used the L_2 norm as the regularization of the representation coefficient and the representation residual. Secondly, the algorithm applied an iteration method to compute the coefficients and the residual, then constructed different update templates according to dif-

收稿日期: 2017-04-21; 修订日期: 2017-05-12.

基金项目: 国家自然科学基金资助项目 (No. 61401425)

ferent non-zero ratio of the residual which was preprocessed by opening operator, and employed the incremental principal component analysis method to learn new PCA subspace basis and PCA subspace mean. Through this way, the algorithm enforced the subspace basis and the subspace mean to describe the variation of the target continuously and accurately during tracking progress. Finally, experimental results on qualitative and quantitative aspects analysis showed that average center location error of the proposed algorithm was 5.3 pixels in all 10 experimental sequences, and average overlap rate was 77%. Compared with eight state-of-art algorithms, the proposed algorithm obtains a more precise result and has better robustness and can meet tracking requirements in more situations.

Key words: sparse representation; visual tracking; subspace representation; particle filter

1 引言

视觉跟踪技术的研究在近年来已取得了可观的进展^[1-4],并且大量的视觉跟踪技术已被应用于各种场景。然而,跟踪技术对于跟踪过程中遇到的遮挡、尺度变化、光照变化以及背景混杂等问题的鲁棒性仍然有提升的空间,因此,视觉跟踪技术的研究仍然具有重要的意义。

2008年,Ross提出了基于主成分分析(Principal Component Analysis, PCA)子空间的增量视觉跟踪(Incremental Visual Tracking, IVT)算法^[5],该算法在粒子滤波框架下利用 PCA 子空间正交基向量的线性组合构建各个粒子的表示模型。研究表明,IVT 算法对于目标区域的光照变化的处理能力较好,但是其对于遮挡等情况的敏感性较高^[6]。另一方面,稀疏表示理论已被应用于计算机视觉领域的人脸识别^[7]、图像分类^[8]及前景检测^[9]等方向。在^[7]的研究基础上,Xue^[10]在 2009 年首次将稀疏表示理论应用于视觉跟踪中,提出基于 L_1 最小化优化的稀疏表示跟踪算法(L_1 -minimization Tracking, L1T)。L1T 算法通过^[5]中的粒子滤波框架采集每一时刻的候选粒子,然后结合稀疏表示理论构建每个粒子的稀疏表示模型,并利用各粒子的表示重构误差作为选择目标的评分依据。L1T 算法能够取得较高的跟踪精度,但是其求解 L_1 优化问题的次数与算法粒子数量成正比,这使得算法的计算复杂度较高,无法满足跟踪的实时性要求。Xue^[11]在 2011 年提出了基于最小边界误差的稀疏跟踪算法,该算法利用最小边界误差筛选候选粒子,可以减少要求解最优化问题的粒子的数量,从而降低了算法的计算代价。Bao^[12]利用加速近似梯度的方法

求解表示模型系数,一定程度上提高了算法的运行效率。Xiao^[13]利用 L_2 范数近似 L_1 范数约束,获得了较高的近似精度,并且在速度上获得较大的提高。Wang^[14]采用最小软阈值平方回归的方法评估各候选粒子,获得了较高的跟踪精度。Xing^[15]依据目标的变化构建多个目标模板字典,并且在跟踪过程中以不同的更新周期对各个字典进行更新。相比于单一字典,该算法对于跟踪过程中目标的表示更加鲁棒,并且能够更加精确的描述目标的变化。Bai^[16]采用结构化的稀疏表示,获得了鲁棒性更好的表示模型。Shreyamsha Kumar^[17]提出利用 PCA 正交基作为被跟踪目标的表示字典,并且采用权值矩阵对各个粒子的评分进行加权。Wang^[18]采用二维主成分分析的方法,获得了目标基于矩阵形式的表示模型。

本文提出一种基于子空间表示与稀疏表示的联合表示模型跟踪算法。首先,本文算法在 PCA 子空间中利用子空间正交基向量集 U 、空间均值 μ 以及表示残差 e 构建目标的表示模型。然后,算法利用 L_2 范数分别约束表示系数 c 与表示残差 e 的稀疏度。其次,算法采用一种基于循环迭代的方式分步求解优化问题的最优解。最后,算法根据经过开运算处理的表示残差中非零元素的不同比率将更新分为三种情况,并采用增量主成分分析方法在线学习新的正交基与均值。相比于文献^[14],本文算法的改进主要有 3 点:(1)本文表示模型包含表示系数 c 与表示残差 e ,而文献^[14]中的方法没有考虑 c 和 e 的信息;(2)本文算法采用的字典更新分为 3 种情况,可以有效处理跟踪过程中遇到的多种情况,而文献^[14]只采用了一种全更新的方式;(3)本文算法对残差项 e 进行了开运算操作,可以提高计算的精确度。而文献^[14]没有对噪声项进行处理。在实验部

分,本文在10个实验序列测试算法,并将结果与8个主流算法进行对比,从定性与定量两个方面进行分析。

2 相关理论

本文利用6个相互独立的仿射参数描述目标状态^[1],即 $\mathbf{x} = \{r, c, \theta, s, \alpha, \phi\}$, 分别对应于目标的水平位移、垂直位移、角度、尺度、宽高比和扭曲度。在仿射参数描述的基础上,粒子滤波框架通过选取 N 个候选粒子近似目标的后验概率分布,并根据最大后验估计准则选取跟踪目标。假设时刻 t 观测向量为 \mathbf{z}_t , 候选粒子为 $\mathbf{x}_t^i (i=1, 2, \dots, N)$ 。依据准则,可通过公式(1)求出 t 时刻的最优目标状态 $\hat{\mathbf{x}}_t$:

$$\hat{\mathbf{x}}_t = \operatorname{argmax}_{\mathbf{x}_t^i} p(\mathbf{x}_t^i | \mathbf{z}_t), \quad (1)$$

其中 $p(\mathbf{x}_t^i | \mathbf{z}_t)$ 为后验概率,可通过贝叶斯定理求得:

$$p(\mathbf{x}_t^i | \mathbf{z}_t) \propto p(\mathbf{z}_t | \mathbf{x}_t^i) \int p(\mathbf{x}_t^i | \mathbf{x}_{t-1}) p(\mathbf{x}_{t-1} | \mathbf{z}_{t-1}) d\mathbf{x}_{t-1}, \quad (2)$$

式中: $p(\mathbf{x}_{t-1} | \mathbf{z}_{t-1})$ 是 $t-1$ 时刻的跟踪结果, $p(\mathbf{x}_t^i | \mathbf{x}_{t-1})$ 是状态转移模型,假设 $p(\mathbf{x}_t^i | \mathbf{x}_{t-1})$ 满足高斯分布:

$$p(\mathbf{x}_t^i | \mathbf{x}_{t-1}) = N(\mathbf{x}_t^i; \mathbf{x}_{t-1}, \mathbf{\Omega}), \quad (3)$$

其中 $\mathbf{\Omega}$ 是对角矩阵,其元素为 \mathbf{x} 中参数的标准差:

$$\mathbf{\Omega} = \operatorname{diag}(\sigma_r^2, \sigma_c^2, \sigma_\theta^2, \sigma_s^2, \sigma_\alpha^2, \sigma_\phi^2), \quad (4)$$

$p(\mathbf{z}_t | \mathbf{x}_t^i)$ 是观测模型,用于评估粒子与目标的相似程度。在本文中,目标的观测模型正比于粒子的评分。

在 t 时刻根据 $t-1$ 时刻的结果与状态转移模型抽取 n 个候选粒子,并根据状态 \mathbf{x}_t^i 从图像中抽取观测图像 $\mathbf{y}_t^i \in \mathbf{R}^{d \times 1}$, d 表示图像列向量化后的维度。通过这种方式,可得到时刻 t 的粒子集 $\{\mathbf{x}_t^i\}_{i=1}^n$ 以及观测图像集 $\mathbf{Y} = [\mathbf{y}^1, \mathbf{y}^2, \dots, \mathbf{y}^n] \in \mathbf{R}^{d \times n}$ 。

3 本文算法

3.1 算法模型

在得到候选粒子集合 \mathbf{Y} 后。本文算法利用 PCA 子空间的正交基向量集 \mathbf{U} 、空间均值 $\boldsymbol{\mu}$ 以及

残差 \mathbf{e} 表示粒子:

$$\mathbf{y}^i = \mathbf{U} \cdot \mathbf{c}^i + \boldsymbol{\mu} + \mathbf{e}^i, \quad (5)$$

其中: $\mathbf{y}^i \in \mathbf{R}^{d \times 1}$ 是粒子 i 的观测值, $\mathbf{U} \in \mathbf{R}^{d \times m}$ 是 PCA 子空间的正交基向量集。 m 是 \mathbf{U} 中正交向量的数量。 $\boldsymbol{\mu} \in \mathbf{R}^{d \times 1}$ 是在跟踪初始化时算法根据目标真值抽取的图像向量。本文算法通过增量主成分分析^[3]的方法更新 \mathbf{U} 和 $\boldsymbol{\mu}$ 。 $\mathbf{c}^i \in \mathbf{R}^{m \times 1}$ 是 \mathbf{y}^i 在 \mathbf{U} 上的表示系数。 $\mathbf{e}^i \in \mathbf{R}^{d \times 1}$ 是表示残差。图1中给出了目标利用 PCA 子空间基向量集、子空间空间均值以及残差进行表示的表示模型说明。

图1 表示模型说明

Fig. 1 Illustration of representation model

图1为 Dudek 实验序列的数值结果。 \mathbf{y} 为某一时刻的观测向量,正交基向量集 \mathbf{U} 是 $1\ 024 \times 5$ 的矩阵, \mathbf{c} 是 \mathbf{y} 利用 \mathbf{U} 进行表示的系数,此时为 5×1 的列向量。 \mathbf{c} 中的非零元素值表示正交基向量集 \mathbf{U} 中相应模板的权重。 $\boldsymbol{\mu}$ 为 $1\ 024 \times 1$ 的列向量,是此时刻 PCA 子空间的均值。 \mathbf{e} 为 $1\ 024 \times 1$ 的列向量,主要描述表示的残差。算法利用 L_2 范数分别约束 \mathbf{c} 和 \mathbf{e} 的稀疏度:

$$\mathbf{y}^i = \mathbf{U} \cdot \mathbf{c}^i + \boldsymbol{\mu} + \mathbf{e}^i \quad \text{s. t.} \quad \|\mathbf{c}^i\|_2 < \varepsilon, \quad \|\mathbf{e}^i\|_2 < \xi, \quad (6)$$

其中 ε 和 ξ 分别表示 \mathbf{z} 和 \mathbf{e} 的稀疏度,公式写成优化函数形式为:

$$[\hat{\mathbf{c}}^i, \hat{\mathbf{e}}^i] = \min_{\mathbf{c}, \mathbf{e}} \frac{1}{2} \|\tilde{\mathbf{y}}^i - \mathbf{U} \cdot \mathbf{c} - \mathbf{e}\|_2 + \lambda \|\mathbf{c}\|_2 + \beta \|\mathbf{e}\|_2, \quad (7)$$

式中 $\tilde{\mathbf{y}}^i = \mathbf{y}^i - \boldsymbol{\mu}$ 。优化问题(7)是非确定多项式难题,无法直接求得 \mathbf{c} 与 \mathbf{e} 。因此本文采用循环迭代分步求解的方法。

3.2 优化问题求解

优化问题(7)无法直接求解 \mathbf{c} 和 \mathbf{e} 。若是固定 \mathbf{c} 或 \mathbf{e} 其中一个,另一个就可直接求解。固定 \mathbf{e} , 优化问题就变为式(8)形式:

$$\hat{\mathbf{c}}^i = \min_{\mathbf{c}} \frac{1}{2} \|(\tilde{\mathbf{y}}^i - \hat{\mathbf{e}}^i) - \mathbf{U} \mathbf{c}\|_2 + \lambda \|\mathbf{c}\|_2, \quad (8)$$

这是最小二乘问题,并且 \mathbf{U} 具有正交性,因此, \mathbf{c} 可以直接通过公式(9)直接求解:

$$\hat{\mathbf{c}}^i = (\mathbf{U}^T \mathbf{U})^{-1} \mathbf{U}^T \frac{\tilde{\mathbf{y}}^i - \mathbf{e}^i}{1+\lambda} = \mathbf{U}^T \frac{\tilde{\mathbf{y}}^i - \mathbf{e}^i}{1+\lambda}, \quad (9)$$

此时,将求解的最优值 $\hat{\mathbf{c}}^i$ 作为 \mathbf{c} 的固定值,优化问题就转换为:

$$\min_{\mathbf{e}} \frac{1}{2} \| (\tilde{\mathbf{y}}^i - \mathbf{U}\hat{\mathbf{c}}^i) - \mathbf{e}^i \|_F^2 + \beta \| \mathbf{e}^i \|_2, \quad (10)$$

可以得到:

$$\hat{\mathbf{e}}^i = \frac{\tilde{\mathbf{y}}^i - \mathbf{U}\hat{\mathbf{c}}^i}{1+\beta}, \quad (11)$$

通过上述两步循环执行,就可以得到优化问题(7)的最优解 $\hat{\mathbf{c}}^i$ 和 $\hat{\mathbf{e}}^i$ 。为了平衡速度与精度,算法对循环次数设定了停止条件,当循环次数达到最大规定次数,或者优化函数 L 的两次迭代间的函数值下降小于阈值时,停止迭代。表 1 中给出了分步循环迭代的过程的整体流程说明。

表 1 分步循环迭代方法

Tab.1 Summary of the Iterative Method

输入: 粒子观测向量 $\tilde{\mathbf{y}}^i$, 正交基向量集 \mathbf{U} , 正则化参数 λ 和 β
初始化: $j=1, \mathbf{e}_{(j)}^i=0$,
循环: 通过 $\mathbf{c}_{(j+1)}^i = \mathbf{U}^T \frac{\tilde{\mathbf{y}}^i - \mathbf{e}_{(j)}^i}{1+\lambda}$ 计算 $\mathbf{c}_{(j+1)}^i$
通过 $\mathbf{e}_{(j+1)}^i = \frac{\tilde{\mathbf{y}}^i - \mathbf{U}\mathbf{c}_{(j+1)}^i}{1+\beta}$ 计算 $\mathbf{e}_{(j+1)}^i$
$j \leftarrow j+1$
直到满足停止条件, 输出 $\hat{\mathbf{c}}^i, \hat{\mathbf{e}}^i$

3.3 观测模型与目标选择

当获得表示系数 \mathbf{c} 与误差向量 \mathbf{e} 后,算法通过观测模型计算粒子评分,并根据评分计算粒子的观测概率。然后算法选取具有最大观测概率的粒子 \hat{i} 作为当前时刻的跟踪目标。粒子的评分函数为:

$$S_i = \frac{1}{2} \| \tilde{\mathbf{y}}^i - \mathbf{U}\hat{\mathbf{c}}^i - \hat{\mathbf{e}}^i \|_F^2 + \lambda \| \hat{\mathbf{e}}^i \|_2, \quad (12)$$

则具有最大观测概率的粒子 \hat{i} 为当前时刻 t 的跟踪: $\hat{i} = \arg \max_i (p(\mathbf{z}_t | \mathbf{x}_t^i)) = \arg \max_i (\exp(-S_i))$ 。

$$(13)$$

3.4 字典更新

当粒子 \hat{i} 被选定为目标,首先根据 $\mathbf{x}_t^{\hat{i}}$ 从第 t 帧图像抽取选定粒子相应图像 $\mathbf{y}_t^{\hat{i}}$ 。然后利用 $3 \times$

3 大小模板 \mathbf{se} 对表示残差 $\mathbf{e}_t^{\hat{i}}$ 进行开运算处理, \mathbf{se} 模板为:

$$\mathbf{se} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}. \quad (14)$$

经过处理后,算法计算表示残差 $\mathbf{e}_t^{\hat{i}}$ 中非零元素的比率 Ratio,并将其与两个预先设定的阈值 Low 和 High 做比较。根据比较结果更新分为 3 种情况:

(1) Ratio < Low, 表明目标没有出现较大的变化且没有出现被遮挡的情况,因此,算法直接将 $\mathbf{y}_t^{\hat{i}}$ 作为跟踪结果收集到数据集 \mathbf{Q} ;

(2) Low < Ratio < High, 表示目标出现部分遮挡的。算法采用一种部分更新的方法,利用 $\mathbf{e}_t^{\hat{i}}$ 中的零元素与非零元素位置作为指导,重新构建待更新的数据 $\bar{\mathbf{y}}$, 然后将其收集到数据集 \mathbf{Q} 中:

$$\bar{\mathbf{y}}(j) = \begin{cases} \mathbf{y}_t^{\hat{i}}(j) & \text{if } \mathbf{e}_t^{\hat{i}}(j) = 0 \\ \boldsymbol{\mu}(j) & \text{if } \mathbf{e}_t^{\hat{i}}(j) \neq 0 \end{cases} \\ j = (1, 2, \dots, d). \quad (15)$$

(3) Ratio > High, 表示目标被遮挡过大,算法此时选择不收集跟踪结果。

当收集一定数量的待更新数据以后,算法采用增量主成分分析方法^[5]从数据集 \mathbf{Q} 中学习新的正交基向量集 $\bar{\mathbf{U}}$ 和空间均值 $\bar{\boldsymbol{\mu}}$ 。然后读取下一帧跟踪图像,将正交基向量集 $\bar{\mathbf{U}}$ 和空间均值 $\bar{\boldsymbol{\mu}}$ 定义为新一帧跟踪的基向量集 \mathbf{U} 、空间均值 $\boldsymbol{\mu}$ 。

4 实验结果

本文算法运行平台: CPU 为 i3-3220, 内存 4 G, 编译环境为 MATLAB2013a, 操作系统为 Win7。跟踪算法的参数设置为: 抽取的图像模板大小均归一化为 32×32 pixel 大小, 在处理过程中, 图像均列向量化为大小为 $d \times 1$ 的向量, 且 $d = 1024$ 。 \mathbf{U} 中的模板数量 $m = 11$ 。 粒子数量 $n = 600$ 。 每 5 帧更新一次 \mathbf{U} 和空间中心 $\boldsymbol{\mu}$ 。 λ 为 0.3, β 为 0.4, 最大循环迭代次数为 5 次, 下降阈值为 0.001。 本文在 10 个实验序列上测试算法, 并将结果与 8 个现今主流算法作对比(分别为 MTT^[19], L1APG^[12], SCM^[20], IVT^[5], TLD^[21], MIL^[22], CT^[23] 和 CSK^[24])。 然后从定性与定量两个方面对实验结果进行分析。 为公平起见, 对比算法源码均为作者提供代码, 并且在跟踪首帧

均以相同初始位置标记目标。图 2 为 9 种算法在 10 个实验序列上的部分时刻跟踪结果。

4.1 计算复杂度分析

本文提出的算法结合了子空间表示与稀疏表示理论。L1APG 算法是经典的稀疏跟踪算法,其

求解优化问题的计算复杂度为 $O(k(dm+d^2))$ 。其中 k 为求解优化的迭代次数, m 为字典中模板的数量。而本文所采用的求解方法的计算复杂度为 $O(k(dm+d))$ 。相比于原 L1APG 跟踪算法,本文提出的算法在速度上获得了较大地提升。

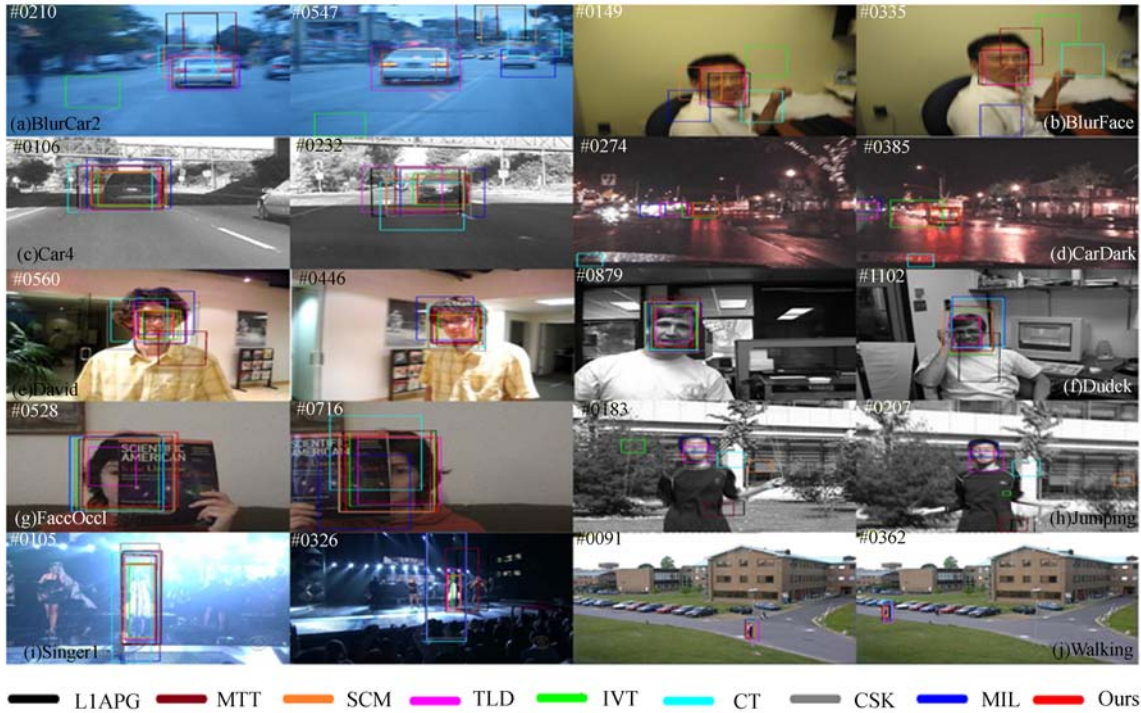


图 2 9 种跟踪算法在 10 个实验序列上的部分跟踪结果图

Fig. 2 Part of tracking results of nine methods in ten image sequences

4.2 定性分析

在图 1 中给出了 9 种实验算法在 10 个实验序列上的部分跟踪结果图。TLD 在对于运动模糊问题的处理能力较好(如 BlurCar2、BlurFace 以及 Jumping),这是因为其采用学习-检测的机制,能够准确学习目标的外观特征并用其作为检测的依据。但是由于 TLD 的学习机制是基于灰度特征,并没有考虑尺度变化,因此,其对于尺度变化、光照变化以及背景混杂的鲁棒性较低(如 Car4、Singer1 以及 CarDark)。CSK 将核函数引入 TLD 的学习机制中,提高了其对于尺度变化与背景混杂问题的鲁棒性,然而这也导致其对于光照变化的处理能力有所下降,在光照变化的处理能力上低于 TLD。MIL 算法是基于多示例学习的算法,其采用判别模型作为粒子的评估依据,当目标区域出现明显的模糊现象时, MIL 算法的分类器无法准确区分目标与背景,这导致算法对于运

动模糊问题的鲁棒性较低(BlurCar2, BlurFace 以及 Jumping)。CT 算法通过稀疏观测矩阵对特征进行编码,然后利用分类器选择目标, CT 算法的分类器对于运动模糊等问题的鲁棒性较低,这也是因为目标区域出现模糊现象时,目标的细节信息丢失较严重,导致分类器无法获得足够的信息,从而导致分类器对于目标与背景的区分能力有所下降。IVT 跟踪算法是基于子空间的跟踪算法,其对于光照、姿态以及尺度变化的处理能力较高(Car4、David 以及 Walking),但是其对于运动模糊以及遮挡等问题的处理能力较差(FaceOccl 以及 BlurCar2),这是因为子空间表示对于遮挡问题较为敏感,并且在目标出现运动模糊等现象时, IVT 算法受到细节信息丢失的影响,导致目标的评估精度下降,从而出现漂移问题。L1APG、MTT 以及 SCM 算法是基于稀疏表示的跟踪算法,其能够获得目标精确鲁棒的表示。并且算法

采用粒子滤波跟踪框架作为目标的运动描述模型,能够处理程度较轻的运动模糊问题。但是 L1APG 算法仅采用表示重构误差作为粒子的评估依据,这导致在目标出现剧烈外观变化时(Singer1, Car4),L1APG 无法精确计算各个粒子的评分,出现漂移甚至丢失目标的问题。MTT 算法将粒子定义为单一任务,采用多任务机制可以获得粒子之间结构关系,提升了稀疏表示跟踪算法对于光照问题的处理能力(Singer1, Car4),但是其对于遮挡问题的处理能力有所下降(FaceOcc1)。SCM 算法结合生成表示模型与判别表示模型,并通过综合两个模型评估结果作为目标的选择依据。判别模型的评分容易受到细节信息丢失的影响。生成模型易受到影响出现偏差。通过结合两个模型,SCM 算法的适应能力较强,能够在多个序列上获得较高的跟踪精度。本文算法是将稀疏表示约束引入子空间表示跟踪算法中,结合了稀疏表示与子空间表示的优势,能够有效处理光照、遮挡以及背景混杂等问题。并且算法利用粒子滤波描述目标的图像帧间的运动,能够处理快速运动等变化。在粒子评估部分,本文算法构建了包含表示残差的粒子评估模型,在目标出现模糊现象时能够精确计算评估各个粒子。在最后,算法采用增量主成分分析的方法在线更新子空间正交基向量集 \bar{U} 以及空间均值 $\bar{\mu}$,

这使得算法能够准确描述目标的外观变化,相比于其他算法,本文算法能够适应更加广泛的跟踪场景。

4.3 定量分析

本文分别采用平均中心误差(Average Centre Location Error, ACLE)和平均覆盖率(Average Overlap Rate, AOR)评估算法在实验序列上的定量评价结果。假设算法标定目标区域 \mathbf{R}_T 和目标的真值 \mathbf{R}_G ,每个时刻的覆盖率结果为^[25]:

$$Rate = \frac{\text{area}(\mathbf{R}_T \cap \mathbf{R}_G)}{\text{area}(\mathbf{R}_T \cup \mathbf{R}_G)}, \quad (15)$$

假设算法标定目标位置为 l_x 和 l_y 。目标的真值位置为 g_x 和 g_y ,那么每一时刻的目标的中心误差计算为公式为:

$$CLE = \sqrt{(l_x - g_x)^2 + (l_y - g_y)^2}. \quad (16)$$

当计算所有时刻的中心误差和覆盖率以后,对结果取平均值得到平均中心误差(ACLE)与平均覆盖率(AOR)。表 2 中给出了所有算法在实验序列上的平均中心位置误差与覆盖率的定量分析结果,其中最优算法用红色文字进行标示,次优算法利用蓝色文字进行标示。图 3 中给出了所有算法的逐帧中心误差曲线,各个子图中,横坐标为实验序列的帧序号,纵坐标单位为像素(彩图见期刊电子版)。

表 2 平均中心位置误差与平均覆盖率

Tab. 2 Average center location error and average overlap rate (unit: Pixel/%)

ACLE/AOR	L1APG	MTT	SCM	TLD	IVT	CT	CSK	MIL	Ours(Sparse)
BlurCar2	144/12.2	141/11.2	125/19.0	6.5/73.0	155/14.0	116/23.2	6.0/75.8	114/25.0	4.5/87.3
BlurFace	17.1/73.4	80.8/33.8	108/18.4	3.7/88.1	157/13.8	120/15.1	9.7/77.8	72.0/27.6	6.7/84.3
Car4	77.0/24.9	22.3/44.7	4.3/75.7	12.8/63.2	2.1/87.5	86.0/21.3	19.1/46.6	50.8/25.8	1.7/90.4
Cardark	1.0/88.4	1.6/82.6	1.3/84.4	27.5/44.9	8.4/66.3	119/30.2	3.2/75.6	43.5/19.6	1.3/86.4
FaceOcc1	13.9/75.2	33.1/70.2	4.3/79.4	5.1/58.5	4.8/72.6	10.4/63.7	17.6/76.5	16.8/56.6	3.8/77.3
David	23.4/53.5	14.1/29.0	10.7/72.1	18.0/71.8	9.6/64.4	26.5/49.5	13.3/40.1	17.6/42.8	9.2/63.2
Dudek	17.3/69.1	21.0/75.8	13.0/76.8	27.3/64.7	18.4/75.2	25.8/64.6	11.9/71.6	29.9/70.7	15.5/69.2
Jumping	83.7/15.0	84.6/15.5	65.9/12.2	5.94/66.3	61.6/12.4	47.7/10.2	85.9/10.1	9.9/52.1	4.7/66.4
Singer1	53.3/28.5	36.2/34.0	2.7/86.8	7.9/72.5	11.3/57.4	15.5/34.8	14.0/35.9	16.3/35.6	3.8/79.8
Walking	3.3/75.3	3.5/66.6	2.5/71.1	10.2/44.6	1.6/76.6	6.9/52.0	7.1/53.7	3.4/54.6	1.9/70.0
Average	43.4/51.6	43.8/46.3	33.8/59.6	12.5/64.8	43.0/54.0	57.4/36.5	18.8/56.4	37.4/41.0	5.3/77.4

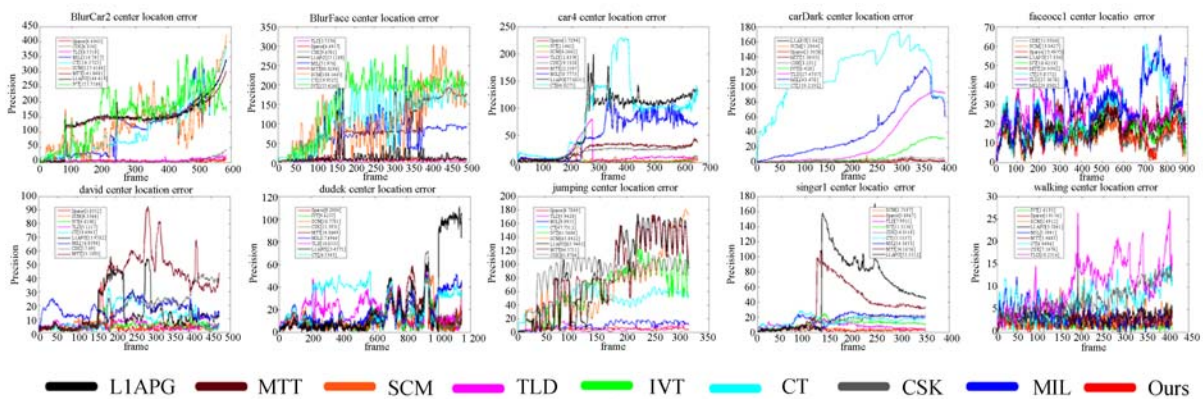


图 3 九种跟踪算法在 10 个实验序列上的逐帧中心误差曲线(横坐标单位:帧序,纵坐标单位:像素)

Fig. 3 Center location error of nine methods in ten image sequences by frames

5 结 论

本文提出的算法结合了子空间表示与稀疏表示的优势,首先,通过在 PCA 子空间对目标进行表示,并且利用 L_2 范数约束表示系数与表示残差,这使得算法能够有效处理跟踪过程中的光照、尺度、遮挡以及背景混杂等问题。其次,算法采用粒子滤波框架作为运动模型,能够有效的描述目标帧间的运动,尤其是快速运动。在粒子评估部分,算法构建的粒子评估模型包含表示残差,这使得算法在目标出现运动模糊的情况下仍然能够获得较高的评估精度,从而准确的选择当前时刻的

目标。然后在更新部分,算法通过开运算对表示残差进行处理,滤除表示残差中的噪声信息,接下来根据处理后的表示残差中的非零元素比率与预先设定阈值的比较结果,将更新方式分为 3 种情况,这使得算法能够灵活准确更新 U 和 μ ,使得算法准确描述目标的变化,完成持久稳定的跟踪。最后,本文在 10 个实验序列上对算法进行了测试,并将结果与 8 个现今主流算法进行对比。根据实验结果可知,本文算法的平均中心误差为 5.3,平均覆盖率为 77%,优于其他跟踪算法。因此,相比于其他对比算法,本文算法能够获得较高的跟踪精度,并且能够适应更加广泛的跟踪场景需求。

参考文献:

- [1] 王春平,王峰,刘江义,等. 基于色度饱和度-角度梯度直方图特征的尺度自适应核相关滤波跟踪[J]. 光学精密工程, 2016, 24(9): 2293-2301.
WANG CH P, WANG W, LIU J Y, *et al.*. Scale adaptive kernelized correlation filter tracking based on HHS-OG feature [J]. *Opt. Precision Eng.*, 2016, 24(9): 2293-2301. (in Chinese)
- [2] 王峰,王春平,李军,等. 特征融合和模型自适应更新相结合的相关滤波目标跟踪[J]. 光学精密工程, 2016, 24(8): 2059-2066.
WANG W, WANG CH P, LI J, *et al.*. Correlation filter tracking based on feature fusing and model adaptive updating [J]. *Opt. Precision Eng.*, 2016, 24(8): 2059-2066. (in Chinese)
- [3] 毛建森,屈玉福. 基于压缩感知的变尺度目标跟踪技术[J]. 液晶与显示, 2016, 31(5): 497-505.
MAO J S, QU Y F. Tracking of variable scale object based on compressive sensing [J]. *Chinese Journal of Liquid Crystals and Displays*, 2016, 31(5): 497-505. (in Chinese)
- [4] 孙保基,张葆,宋策,等. 基于角点增强改进的 TLD 目标跟踪算法[J]. 液晶与显示, 2016, 31(9): 921-928.
SUN B J, ZHANG B, SONG C, *et al.*. Improved TLD object tracking algorithm based on corner reinforced[J]. *Chinese Journal of Liquid Crystals and Displays*, 2016, 31(9): 921-928. (in Chinese)
- [5] ROSS D A, LIM J, LIN R S, *et al.*. Incremental learn-

- ing for robust visual tracking[J]. *International Journal of Computer Vision*, 2008, 77(1-3): 125-141.
- [6] LIU R SH, BAI SH SH, SU ZH X, *et al.*. Robust visual tracking via L0 regularized local low-rank feature learning[J]. *Journal of Electronic Imaging*, 2015, 24(3): 033012.
- [7] WRIGHT J, YANG A Y, GANESH A, *et al.*. Robust face recognition via sparse representation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009, 31(2): 210-227.
- [8] 唐中奇, 付光远, 陈进, 等. 基于多尺度分割的高光谱图像稀疏表示与分类[J]. *光学 精密工程*, 2015, 23(9): 2708-2714.
- TANG ZH Q, FU G Y, CHEN J, *et al.*. Multi-scale segmentation-based sparse coding for hyperspectral image classification [J]. *Opt. Precision Eng.*, 2015, 23(9): 2708-2714. (in Chinese)
- [9] 隋中山, 李俊山, 张斌, 等. 张量低秩表示和时空稀疏分解的视频前景检测[J]. *光学 精密工程*, 2017, 25(2): 529-536.
- SUI ZH SH, LI J SH, ZHANG J, *et al.*. Video foreground detection of tensor low-rank representation and spatial-temporal sparsity decomposition[J]. *Opt. Precision Eng.*, 2017, 25(2): 529-536. (in Chinese)
- [10] MEI X, LING H B. Robust visual tracking using l_1 minimization[C]. *Proceedings of the 12th International Conference on Computer Vision*, IEEE, 2009: 1436-1443.
- [11] MEI X, LING H B, WU Y, *et al.*. Minimum error bounded efficient l_1 tracker with occlusion detection[C]. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, 2011: 1257-1264.
- [12] BAO C L, WU Y, LING H B, *et al.*. Real time robust L1 tracker using accelerated proximal gradient approach[C]. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, 2012: 1830-1837.
- [13] XIAO Z Y, LU H CH, WANG D. L2-RLS-based object tracking[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2014, 24(8): 1301-1309.
- [14] WANG D, LU H CH, YANG M H. Least soft-threshold squares tracking [C]. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, 2013: 2371-2378.
- [15] XING J L, GAO J, LI B, *et al.*. Robust object tracking with online multi-lifespan dictionary learning [C]. *Proceedings of IEEE International Conference on Computer Vision*, IEEE, 2013: 665-672.
- [16] BAI T X, LI Y F. Robust visual tracking with structured sparse representation appearance model [J]. *Pattern Recognition*, 2012, 45(6): 2390-2404.
- [17] KUMAR B K S, SWAMY M N S, AHMAD M O. Weighted residual minimization in PCA subspace for visual tracking [C]. *Proceedings of IEEE International Symposium on Circuits and Systems*, IEEE, 2016: 986-989.
- [18] WANG D, LU H CH. Object Tracking via 2DPCA and l_1 -regularization [J]. *IEEE Signal Processing Letters*, 2012, 19(11): 711-714.
- [19] ZHANG T ZH, GHANEM B, LIU S, *et al.*. Robust visual tracking via multi-task sparse learning [C]. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2012: 2042-2049.
- [20] ZHONG W, LU H CH, YANG M H. Robust object tracking via sparse collaborative appearance model[J]. *IEEE Transactions on Image Processing*, 2014, 23(5): 2356-2368.
- [21] KALAL Z, MIKOLAJCZYK K, MATAS J. Tracking-learning-detection [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34(7): 1409-1422.
- [22] BABENKO B, YANG M H, BELONGIE S. Robust object tracking with online multiple instance learning[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, 33(8): 1619-1632.
- [23] ZHANG K H, ZHANG L, YANG M H. Real-

time compressive tracking[C]. *Proceedings of the 12th European Conference on Computer Vision*, Springer, 2012: 864-877.

- [24] 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, Springer, 2012: 702-715.

- [25] EVERINGHAM M, VAN GOOL L, WILLIAMS C K I, *et al.*. The Pascal visual object classes (VOC) challenge[J]. *International Journal of Computer Vision*, 2010, 88(2): 303-338.

作者简介:



陈典兵(1990—),男,吉林长春人,博士研究生,2012年于吉林大学获得学士学位,主要从数字图像处理、稀疏表示、目标跟踪方面的研究。E-mail: chendi-anbing1934@163.com

导师简介:



朱明(1964—),男,江西南昌人,研究员,博士生导师,1985于南京航空学院年获得学士学位,1991年于中国科学院长春光学精密机械与物理研究所获得硕士学位,主要从事数字图像处理、电视跟踪、自动目标识别等方面的研究。E-mail: zhu_mingca@163.com