

一种结合空间信息和稀疏字典优化的目标跟踪算法

胡秀华, 郭雷, 李晖晖, 鹿馨

(西北工业大学自动化学院, 西安 710129)

摘要: 针对复杂场景中目标表观变化引起的跟踪漂移问题, 提出一种新的基于稀疏表示的目标跟踪算法. 该算法通过稀疏性和空间相关性正则约束得到一种优化的目标代价函数, 利用拉格朗日对偶理论和加速近端梯度方法完成字典优化, 并利用最大池化理论和空间金字塔方法得到降维的且包含更多空间信息的目标模板系数和候选样本系数. 实验结果表明, 所提出的算法在背景干扰、光照变化、形变、运动模糊、严重遮挡等多种复杂场景中都能取得较为鲁棒的跟踪效果.

关键词: 目标跟踪; 稀疏表示; 空间信息; 字典优化

中图分类号: TP391

文献标志码: A

An object tracking algorithm combining spatial information and sparse dictionary optimization

HU Xiu-hua, GUO Lei, LI Hui-hui, LU Xin

(College of Automation, Northwestern Polytechnical University, Xi'an 710129, China. Correspondent: HU Xiu-hua, E-mail: huxh@mail.nwpu.edu.cn)

Abstract: Aiming at the problem of tracking drift caused by object appearance change in complex scene, a novel object tracking algorithm based on sparse representation is proposed. An optimized objective cost function is designed with the sparsity and spatial correlation regularization constraint. The Lagrange dual theory and the accelerate proximal gradient approach are used to complete the dictionary optimization. By using the maximum pooling theory and the spatial pyramid method, the coefficients of the object template and candidate samples with the reduced dimension and more spatial information are obtained. Experimental results show that the proposed algorithm can perform robust tracking effect in a variety of complex scene, such as background clutters, illumination variation, deformation, motion blur, heavy occlusion, and so on.

Keywords: object tracking; sparse representation; spatial information; dictionary optimization

0 引言

目标跟踪作为计算机视觉中的一个重要研究领域, 不仅可以提供目标运动状态和轨迹, 还可以为运动检测、行为分析、场景理解等提供重要数据, 广泛应用于智能监控、人机交互、视觉导航等方面^[1-2]. 但是, 由于受到目标自身姿态、形状变化等内部因素, 光照变化、运动模糊、杂波干扰、旋转、遮挡等外部因素的影响, 目标表观模型动态变化, 研究如何进行运动目标的准确跟踪仍然是一项具有挑战性的问题.

近年来, 基于稀疏表示理论的目标跟踪算法取得了一定的研究进展. 文献[3]通过引入微小模板解决了目标变形和遮挡影响等问题, 但目标模板数量增多,

算法复杂度高. 针对 L_1 算法跟踪效率低的问题, 文献[4]改进了琐碎模板, 并运用加速近端梯度方法完成 L_1 范数最小化求解, 但所提出的整体稀疏表示模型不适合处理遮挡和目标形变问题. 为解决目标表观变化问题, 文献[5]给出了一种结构局部稀疏表示模型, 并利用增量空间学习和稀疏表示相结合的方式对模板进行更新, 但是跟踪性能容易受运动模型的影响. 文献[6]采用结合原始目标模板与当前帧粒子模板的联合目标函数, 并迭代更新原始目标模板, 适用于解决场景中含有与目标模板相似背景时的跟踪问题. 文献[7]提出采用判别式模型与生成式模型联合的稀疏表示方法, 对复杂场景的跟踪效果良好, 但是实时性

收稿日期: 2015-12-02; **修回日期:** 2016-02-23.

基金项目: 武器装备预先研究基金项目(51306030205); 航空科学基金项目(20131953022).

作者简介: 胡秀华(1988—), 女, 博士生, 从事目标跟踪、多传感器信息融合的研究; 郭雷(1956—), 男, 教授, 博士生导师, 从事图像处理、模式识别与智能系统等研究.

不高,且没有考虑空间相关性.文献[8]通过构建多任务逆稀疏表示模型,得到判别式稀疏相似性置信图,其中正、负模板的动态更新和加性池化方法的运用可以有效地从复杂背景中提取目标信息,但在处理目标遮挡和运动模糊问题时性能减弱.现有跟踪算法虽然在目标观信息描述中考虑到了增加背景信息、模板更新等改进方法,但是很少有文献在跟踪过程中进行模板字典的优化,以及考虑目标空间几何分布等问题.

为了更好地解决目标在具有背景干扰、光照变化、形变、运动模糊、严重遮挡等多种复杂场景中容易出现跟踪漂移的问题,设计一种结合空间结构信息和稀疏字典优化的目标跟踪算法.首先,利用似然估计理论得到更为稀疏的代价函数,同时考虑目标空间结构相关性,引入拉普拉斯正则项,得到一种优化的目标代价函数;然后,利用聚类方法得到初始化的字典,并借助拉格朗日对偶理论和加速近端梯度方法进行字典优化;最后,基于优化的字典,利用最大池化理论和空间金字塔方法,得到降维的且能够包含更多空间信息的目标模板系数和候选样本系数,并利用相似性判定准则得到与目标模板最相似的候选样本,记为当前目标位置,实现目标的准确跟踪.

1 目标跟踪算法框架

目标跟踪问题可以看作目标运动状态的推理问题.基于贝叶斯理论框架,在第 k 帧,定义目标状态变量 \mathbf{s}_k ,观测集 $\mathcal{O}_k = [\mathbf{o}_1, \mathbf{o}_2, \dots, \mathbf{o}_k]$,有

$$p(\mathbf{s}_k | \mathcal{O}_k) \propto p(\mathbf{o}_k | \mathbf{s}_k) \int p(\mathbf{s}_k | \mathbf{s}_{k-1}) p(\mathbf{s}_{k-1} | \mathcal{O}_{k-1}) d\mathbf{s}_{k-1}. \quad (1)$$

其中: $p(\mathbf{s}_k | \mathbf{s}_{k-1})$ 为目标在连续两帧之间的运动模型; $p(\mathbf{o}_k | \mathbf{s}_k)$ 为在状态 \mathbf{s}_k 下对观测量 \mathbf{o}_k 进行估计的观测模型, $k = 1, 2, \dots$. 目标在第 k 帧的最佳状态可以通过最大后验概率估计获得,即目标运动状态可表示为

$$\hat{\mathbf{s}}_k = \arg \max_{\mathbf{s}_k^\tau} p(\mathbf{s}_k^\tau | \mathcal{O}_k), \quad (2)$$

其中 \mathbf{s}_k^τ 为第 τ 个候选样本的状态变量, $\tau = 1, 2, \dots, \Gamma$, Γ 为候选样本个数.

1.1 目标运动模型

采用仿射变换的 6 个参数描述目标连续两帧之间的旋转、尺度和平移等运动.在第 k 帧,定义目标状态变量 $\mathbf{s}_k = [x_k, y_k, \theta_k, s_k, \eta_k, \phi_k]$,其中 $x_k, y_k, \theta_k, s_k, \eta_k, \phi_k$ 分别为目标在水平与垂直方向上的平移量、旋转角、尺度变化量、高宽比和斜切角. 6 个仿射参数相互独立,采用随机游走模型描述连续两帧之间目标状态的转换,即

$$p(\mathbf{s}_k | \mathbf{s}_{k-1}) = N(\mathbf{s}_k; \mathbf{s}_{k-1}, \Sigma). \quad (3)$$

其中: $\Sigma = \text{diag}\{\sigma_x^2, \sigma_y^2, \sigma_\theta^2, \sigma_s^2, \sigma_\eta^2, \sigma_\phi^2\}$, 对角线上元素

依次为各仿射参数的方差.

1.2 目标观测模型

为了从运动模型得到的 Γ 个候选样本中找到最佳候选样本,在第 k 帧,设定观测模型 $p(\mathbf{o}_k | \mathbf{s}_k)$ 为

$$p(\mathbf{o}_k | \mathbf{s}_k) = L_k^\tau, \quad (4)$$

其中 L_k^τ 为第 τ 个候选样本与目标模板的相似性度量, $\tau = 1, 2, \dots, \Gamma$.

本文核心为,基于稀疏表示理论设计一种新的算法,完成各候选样本与目标模板的鲁棒描述和相似性度量,求得最佳候选样本,实现稳定的目标跟踪.

2 问题描述

将图像转化为灰度图像,对每个图像序列于第 1 帧中手动标出初始矩形目标区域,并归一化为 $n \times n$ 像素,记为 $\mathbf{I} \in \mathbf{R}^{n \times n}$. 记目标区域中心为 $X(h, v)$, 在目标中心位置周围 $\|X_l - X(h, v)\| \leq r_o$ 圆形区域范围内,通过密集采样提取 L 个大小为 $n \times n$ 像素的目标模板组成模板样本集 \mathcal{T} . 其中: X_l 为第 l 个模板样本的中心位置, r_o 为圆形区域半径, $\mathcal{T} = \{\mathbf{T}_1, \mathbf{T}_2, \dots, \mathbf{T}_L\}$, $\mathbf{T}_l \in \mathbf{R}^{n \times n}$ 为第 l 个目标模板样本, $l = 1, 2, \dots, L$. 利用粒子滤波采样得到 Γ 个候选样本,对每一个候选样本 \mathbf{y} 进行分块得到 N 个图像块,每个候选样本图像块可表示为向量 $\mathbf{y}_i, \mathbf{y}_i \in \mathbf{R}^{G \times 1}, i = 1, 2, \dots, N, G$ 为列向量的大小.

在稀疏表示理论框架下,目标跟踪问题可以视为在给定字典 \mathbf{D} 中利用少量的能有效表征目标信息的模板近似表达候选样本信息 \mathbf{y} ,使得 $\mathbf{y} = \mathbf{D}\alpha + \varepsilon$. 其中: α 为稀疏表示系数向量, ε 为候选样本无法由字典完全表示的部分残差向量. α 的稀疏性可以用 L_0 范数来衡量,但是 L_0 范数的最小化是 NP-hard 问题^[9],当候选样本的表示系数足够稀疏时, L_1 范数最小化等价于 L_0 范数的最小化,因此,上述稀疏表示问题可以表示为

$$\min_{\alpha} \|\alpha\|_1, \text{ s.t. } \|\mathbf{y} - \mathbf{D}\alpha\|_F^2 \leq \zeta. \quad (5)$$

其中 ζ 为一个非常小的常量.该模型表明了稀疏表示的两个约束条件,第 1 项约束了信息表示的稀疏性,第 2 项约束了信息表示的准确性.

式(5)描述的问题等价于 LASSO 问题,表示为

$$\min_{\alpha} \|\mathbf{y} - \mathbf{D}\alpha\|_F^2, \text{ s.t. } \|\alpha\|_1 \leq \varsigma. \quad (6)$$

其中 ς 为一个非常小的常量.定义重构误差 $\varepsilon = \mathbf{y} - \mathbf{D}\alpha$,当重构误差 ε 服从高斯分布时,基于稀疏表示理论,设每个图像块的稀疏表示系数解为 α_i ,则有

$$\alpha_i = \arg \min_{\alpha_i} \|\mathbf{y}_i - \mathbf{D}\alpha_i\|_F^2 + \lambda \|\alpha_i\|_1. \quad (7)$$

其中: \mathbf{D} 为由 J 个目标模板样本图像块得到的字典, $\mathbf{D} \in \mathbf{R}^{G \times J}$, λ 为正则化系数.式(7)是一个稀疏约束

的最小方差估计问题,当重构误差服从高斯分布时,问题的解是最大似然解^[9].在目标跟踪问题中,准确性约束比稀疏性约束更重要,在实际应用中,当目标发生遮挡、形变、光照变化等异常情形时,异常像素的重构误差 ε 不再服从高斯分布假设,因此,需要对传统的稀疏表示模型中的准确性约束项进行改进,构建鲁棒性更好的目标表现模型.

3 结合空间信息和稀疏字典优化的跟踪算法

3.1 算法的提出

传统的基于稀疏表示理论的跟踪方法在表征目标方面表现出了很好的优势,但是有两个缺陷;1) 由于字典的过完备性,稀疏表示系数对局部特征的变化较为敏感,局部特征的异常可能会降低字典基向量在表征目标时的准确性,考虑运用最大似然估计方法优化目标代价函数;2) 目标稀疏表示过程中,局部特征之间的空间相关性信息丢失,影响了目标的鲁棒描述,考虑加入拉普拉斯正则约束,描述目标跟踪中位置分布的重建误差,改进损失函数项.

定义行向量形式的字典 $D = [r_1; r_2; \dots; r_G]$, 图像块 y_i 的重构误差为 $\varepsilon_i = [\varepsilon_{1,i}; \varepsilon_{2,i}; \dots; \varepsilon_{G,i}]$, $\varepsilon_{g,i} = y_{g,i} - r_{g,j}\alpha_{j,i}$, $\varepsilon_{g,i}$ 为第 i 个图像块第 g 个像素的重构误差, $i = 1, 2, \dots, N$, $g = 1, 2, \dots, G$, $j = 1, 2, \dots, J$. 假设 $\varepsilon_{1,i}, \varepsilon_{2,i}, \dots, \varepsilon_{G,i}$ 相互独立且服从相同的概率密度分布函数 $f_\theta(\varepsilon_{g,i})$, θ 是描述概率密度函数特征的参数集,则似然函数为

$$L_\theta(\varepsilon_{1,i}, \varepsilon_{2,i}, \dots, \varepsilon_{G,i}) = \prod_{g=1}^G f_\theta(\varepsilon_{g,i}),$$

令

$$g_\theta(\varepsilon_{g,i}) = -\ln f_\theta(\varepsilon_{g,i}),$$

$$-\ln L_\theta(\varepsilon_{1,i}, \varepsilon_{2,i}, \dots, \varepsilon_{G,i}) = \sum_{g=1}^G g_\theta(\varepsilon_{g,i}),$$

则最大似然估计的目标是最大化似然函数 L_θ 或最小化目标函数 $-\ln L_\theta$.

在考虑稀疏性约束的情况下,图像块 y_i 的稀疏表示系数解可表示为

$$\alpha_i = \arg \min_{\alpha_i} \sum_{g=1}^G g_\theta(y_{g,i} - r_{g,j}\alpha_{j,i}) + \lambda \|\alpha_i\|_1, \quad (8)$$

其中 $\alpha_{j,i}$ 为第 i 个图像块稀疏系数的第 j 个元素.

由文献[9]可知,式(8)近似等价于

$$\alpha_i = \arg \min_{\alpha_i} \frac{1}{2} \|\mathbf{W}_i^{1/2}(\mathbf{y}_i - D\alpha_i)\|_F^2 + \lambda \|\alpha_i\|_1. \quad (9)$$

式(9)中,考虑到重构误差对候选样本图像块鲁棒表示的影响,对于图像块 y_i , 给定对角矩阵 \mathbf{W}_i , 定义 $W_{g,g}^i = 1/(1 + \exp(\kappa\varepsilon_{g,i}))$, 其中 κ 为常量, 满足

$W_{g,g}^i \in [0, 1]$, $W_{g,g}^i$ 为第 i 个图像块中第 g 个像素的权值, 重构误差越大表明图像块信息出现异常而应该分配越小权值.

为了更好地描述局部图像块之间的相关性,降低稀疏表示系数对异常因素的敏感性,引入拉普拉斯正则项,使得相似图像块之间能具有良好的一致性,减小重构误差. 定义相似性矩阵 M , $M \in \mathbf{R}^{N \times N}$, M 中的元素 M_{iq} 表示第 i, q 个图像块之间的相似性. 令 $M_{ii} = 0$, 若 y_i 位于 y_q 的最近邻区域内,则依据直方图交互理论^[10]计算相似性度量

$$M_{iq} = \sum_{d=1}^N \min(H_{id}, H_{qd}), \quad H_{id} = \sum_{i=1}^N |\alpha_{i,d}|.$$

若 y_i 不在 y_q 的最近邻区域内,则 $M_{iq} = 0$.

3.2 算法的求解

在考虑稀疏性约束和空间相关性正则约束情况下,候选样本 y 的鲁棒稀疏表示系数解 α 为

$$\alpha = \arg \min_{\alpha} \frac{1}{2} \|\mathbf{W}^{1/2}(\mathbf{y} - D\alpha)\|_F^2 + \lambda \sum_{i=1}^N \|\alpha_i\|_1 + \frac{\rho}{2} \sum_{q=1}^N \|(\alpha_i - \alpha_q)\|_2^2 M_{iq}. \quad (10)$$

其中: $\mathbf{W} = [\mathbf{W}_1, \mathbf{W}_2, \dots, \mathbf{W}_N]$ 为对角矩阵, \mathbf{W}_i 为第 i 个图像块对应的对角矩阵, α 为候选样本稀疏表示系数矩阵, α_i, α_q 为第 i, q 个图像块的稀疏表示系数向量, ρ 为正则化调整系数.

由式(10)可以得到更具有一般性的带有权重和拉普拉斯正则项的稀疏表示解为

$$\alpha = \arg \min_{\alpha} \frac{1}{2} \|\mathbf{W}^{1/2}(\mathbf{y} - D\alpha)\|_F^2 + \lambda \sum_{i=1}^N \|\alpha_i\|_1 + \rho \text{tr}(\alpha L \alpha^T). \quad (11)$$

其中: $\alpha_i \geq 0$, $L = C - M$ 为拉普拉斯矩阵, $C = \text{diag}(C_1, C_2, \dots, C_N)$, $C_i = \sum_{q=1}^N M_{iq}$.

可以看出,当 $W_{g,g}^i$ 取为 2 且 ρ 取 0 时,式(11)等价于初始的稀疏编码问题.

令 $\hat{D} = \mathbf{W}^{1/2}D$, $\hat{y}_i = \mathbf{W}^{1/2}y_i$, 候选样本 y 的鲁棒稀疏表示系数矩阵

$$\beta = \arg \min_{\beta, \hat{D}} \frac{1}{2} \|\hat{y} - \hat{D}\beta\|_F^2 + \lambda \sum_{i=1}^N \|\beta_i\|_1 + \rho \text{tr}(\beta L \beta^T). \quad (12)$$

其中: $\mathbf{1}^T \beta_i = 1$, $\beta_i \geq 0$, $\|\hat{D}_j\|_2^2 \leq 1$, $\mathbf{1}^T \beta_i = 1$ 保证了平移不变性, \hat{D}_j 为字典列向量, $\hat{D}_j \in \mathbf{R}^{G \times 1}$, $j = 1, 2, \dots, J$, 图像块 y_i 的重构误差为 $\varepsilon_i = \hat{y}_i - \hat{D}\beta_i$.

式(12)中,包含 \hat{D} 和 β 的函数是非凸函数,可以将其拆分为两个子部分,分别进行凸优化求解,即固定 β 求解 \hat{D} , 固定 \hat{D} 求解 β .

对于式 (12), 若固定 β 求解 \hat{D} , 则目标函数转化为平方约束的最小二乘模型, 即

$$\hat{D} = \arg \min_{\hat{D}} \frac{1}{2} \|\hat{y} - \hat{D}\beta\|_F^2. \quad (13)$$

采用拉格朗日对偶思想进行求解^[11], 有

$$L(\hat{D}, \lambda) = \frac{1}{2} \text{tr}((\hat{y} - \hat{D}\beta)^T (\hat{y} - \hat{D}\beta)) + \sum_{j=1}^J \lambda_j (\|\hat{D}_j\|_2^2 - 1), \quad (14)$$

其中 λ_j 为对应第 j 列字典向量的对偶变量, $\lambda_j \geq 0$.

对于式 (14), 由 $\partial L(\hat{D}, \lambda) / \partial \hat{D} = 0$, 可得到 \hat{D} 表达式为 $\hat{D} = \hat{y}\beta^T(\beta\beta^T + 2\mathbf{A})^{-1}$, 代入式 (14), 得到拉格朗日对偶表达式

$$\mathcal{D}(\lambda) = \min_{\hat{D}} L(\hat{D}, \lambda) = \frac{\text{tr}(\hat{y}^T \hat{y} - \hat{y} \beta^T (\hat{\beta})^{-1} (\hat{y} \beta^T)^T - 2\mathbf{A})}{2}. \quad (15)$$

其中: $\mathbf{A} = \text{diag}(\lambda)$, $\hat{\beta} = \beta\beta^T + 2\mathbf{A}$.

利用共轭梯度法进行字典 \hat{D} 的更新求解, 对 $\mathcal{D}(\lambda)$ 求一阶导数和二阶偏导, 可以得到

$$\frac{\partial \mathcal{D}(\lambda)}{\lambda_g} = \|\hat{y}\beta^T(\hat{\beta})^{-1} e_g\|_F^2 - 1, \quad (16)$$

$$\frac{\partial^2 \mathcal{D}(\lambda)}{\lambda_g \lambda_j} = -4((\hat{\beta})^{-1} (\hat{y}\beta^T)^T \hat{y}\beta^T (\hat{\beta})^{-1})_{g,j} (\hat{\beta})_{g,j}^{-1}. \quad (17)$$

其中: λ_g 为对应第 g 行字典向量的对偶变量, e_g 为对应的单位向量, $g = 1, 2, \dots, G$. 求得 $\mathcal{D}(\lambda)$ 关于 \mathbf{A} 的极大值 $\hat{\mathbf{A}}$, 得到优化的字典 \hat{D} 为

$$\hat{D} = \hat{y}\beta^T(\beta\beta^T + 2\hat{\mathbf{A}})^{-1}. \quad (18)$$

对于初始字典的选择, 首先对 L 个模板样本进行分块, 并通过 k 均值聚类方法得到字典 D , $D \in \mathbf{R}^{G \times J}$, G 为聚类中心维数, J 为聚类中心个数, 聚类中心表示最具有代表性的模板样本图像块.

对于式 (12), 固定 \hat{D} 求解 β , 目标函数转化为

$$\beta = \arg \min_{\beta} \frac{1}{2} \|\hat{y} - \hat{D}\beta\|_F^2 + \lambda \sum_{i=1}^N \|\beta_i\|_1 + \rho \text{tr}(\beta L \beta^T). \quad (19)$$

令 $\mathbf{1} \in \mathbf{R}^J$, $\mathbf{1}^* \in \mathbf{R}^N$, $\mathbf{1}$ 、 $\mathbf{1}^*$ 为元素均为 1 的列向量, J 为典型模板样本图像块个数, N 为候选样本图像块个数. 定义

$$\psi(\beta_i) = \begin{cases} 0, & \beta_i \geq 0; \\ +\infty, & \text{otherwise}; \end{cases} \quad i = 1, 2, \dots, N.$$

则式 (19) 可优化为

$$\beta = \arg \min_{\beta} \frac{1}{2} \|\hat{y} - \hat{D}\beta\|_F^2 + \lambda \mathbf{1}^T \beta \mathbf{1}^* + \rho \text{tr}(\beta L \beta^T) + \psi(\beta). \quad (20)$$

利用加速近端梯度方法求解上述最小化问题, 有

$$F(\beta) = \arg \min_{\beta} \frac{1}{2} \|\hat{y} - \hat{D}\beta\|_F^2 + \lambda \mathbf{1}^T \beta \mathbf{1}^* + \rho \text{tr}(\beta L \beta^T), \quad (21)$$

$$G(\beta) = \psi(\beta). \quad (22)$$

其中: $F(\beta)$ 为可微分凸函数, $G(\beta)$ 为非平滑凸函数.

基于加速近端梯度理论, 有

$$\gamma_{t+1} = \arg \min_{\beta} \frac{Q \left\| \beta - \xi_{t+1} + \frac{\nabla F(\xi_{t+1})}{Q} \right\|_F^2}{2} + G(\beta). \quad (23)$$

其中

$$\xi_{t+1} = \gamma_t + \frac{\kappa_{t-1} - 1}{\kappa_t} (\gamma_t - \gamma_{t-1}),$$

$$\kappa_{t+1} = \frac{1 + \sqrt{1 + 4\kappa_t^2}}{2},$$

$$\gamma_0 = \gamma_{-1} = \mathbf{0} \in \mathbf{R}^{J \times N}, \kappa_0 = \kappa_{-1} = 1,$$

Q 为利普希茨常数且 $Q \leq \lambda_{\max}^2 + \rho + 1$, λ_{\max} 为字典 \hat{D} 的最大奇异值, $F(\xi_{t+1})$ 为具有连续梯度的函数, t 为当前迭代时刻.

定义 $g_{t+1} = \xi_{t+1} - \nabla F(\xi_{t+1})/Q$, 由

$$\nabla F(\xi_{t+1}) = -\hat{D}^T (\hat{y} - \hat{D}\xi_{t+1}) + \lambda \mathbf{1} \mathbf{1}^{*T} + \rho \xi_{t+1} (L^T + L)$$

可得

$$g_{t+1} = \xi_{t+1} - \frac{-\hat{D}^T (\hat{y} - \hat{D}\xi_{t+1}) + \lambda \mathbf{1} \mathbf{1}^{*T} + \rho \xi_{t+1} (L^T + L)}{Q}. \quad (24)$$

式 (23) 等价于求解 $\gamma_{t+1} = \max(\mathbf{0}, g_{t+1})$, 由此可得候选样本稀疏系数矩阵 β 为 $\beta = \gamma_{t+1}$, 进而迭代求解式 (23) 和 (24) 得到优化的字典 \hat{D} .

利用优化的字典, 通过最大池化理论和空间金字塔方法^[12]得到降维的且包含更多空间结构信息的目标模板系数向量 φ 和候选样本系数矩阵 ϕ , 实现了对目标模板和候选样本的鲁棒描述. 依据欧氏距离公式 $d(\cdot)$ 判定候选样本系数向量和目标模板系数向量

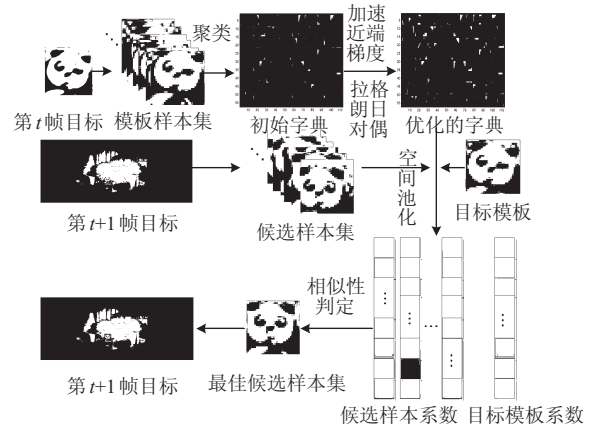


图 1 算法原理

的相似性,得到第 τ 个候选样本与目标模板之间的相似度为 $L^\tau = \min_{\Gamma}(d(\phi^\tau, \varphi))$,其中 ϕ^τ 为第 τ 个候选样本系数向量, $\tau = 1, 2, \dots, \Gamma$. 距离度量值最小的候选样本即为最佳候选样本. 所提出算法原理如图1所示.

3.3 计算复杂度分析

在计算时间复杂度方面,所提出算法耗时主要集中在字典优化和稀疏系数向量的池化部分. 相比于梯度下降算法,在拉格朗日对偶基础上求解优化字典基利用了较小数目的优化变量,每帧计算复杂度取决于对偶变量个数 J ;利用加速近端梯度法求解稀疏系数的计算耗时主要取决于式(24),每帧计算复杂度可表示为 $O(tGN)$, t 为迭代次数, G 为候选样本图像块大小, N 为候选样本图像块个数;稀疏系数向量池化的计算复杂度可表示为 $O(\Gamma(G+P))$, Γ 为候选样本个数, P 为空间金字塔多尺度分块总数.

4 实验分析

实验中算法的仿真环境为Matlab 2012b,计算机处理器为i3-2130,主频为3.40 GHz,内存为4 GB. 手动选择第一帧,归一化目标模板,候选样本大小为 32×32 pixels,选择滑动窗大小为 6×6 pixels,滑动步长为2 pixels. 为了验证本文算法的有效性,选择基准库中的8组具有尺寸变化、光照变化、运动模糊、背景干扰和遮挡等复杂情况的视频序列进行测试,同时与LIAPG^[4]、IVT^[13]、MTT^[14]、LSK^[15]、ASLA^[5]、SCM^[7]、DSSmap^[8]算法进行对比,并采用中心位置误差定量评价各跟踪算法性能. 部分时刻不同算法在每段测试视频序列上的跟踪效果如图2所示.

由图2可见:

animal测试序列中目标做跳跃运动,且存在严重的背景干扰. 在第24帧目标跃起并产生运动模糊,

LSK算法、ASLA算法和本文算法跟踪误差明显增大,但本文算法可以更快地重新跟踪上目标. 在第35帧存在运动模糊和第54帧背景干扰的情况下,由于本文算法考虑了空间相关性信息,可以取得鲁棒的跟踪效果. 整体看本文算法和DSSmap算法、SCM算法都能有效实现稳定跟踪.

bird2测试序列中存在旋转和遮挡等情况. MTT算法在开始便丢失了目标,后续过程中不能重新找回. 目标在第46帧反向运动,各算法跟踪误差都有所增加,但本文算法运用优化的模板字典和空间信息能够快速找准目标,能够有效实现完整跟踪.

car11测试序列目标与背景对比度低,且存在较强光照变化. 当目标分辨率较低时,基于整体模板信息的LIAPG算法和IVT算法性能不稳定,出现不同程度的跟踪漂移. 基于局部信息的LSK算法没能充分利用空间信息,不能很好地将目标从复杂背景中分离出来;而DSSmap算法等都取得了较鲁棒的跟踪效果.

car4测试序列存在目标尺寸变化和背景干扰等情况. 在第186帧受低光照强度和尺寸变化的影响,IVT算法和LSK算法跟踪误差明显增大,并出现持续的跟踪漂移. 在第450帧之后,基于判别式稀疏表示模型的DSSmap算法性能下降,而SCM算法、ASLA算法、LIAPG算法和本文算法取得了较鲁棒的跟踪效果.

faceoccl测试序列存在严重遮挡和外界干扰. 在第216帧受遮挡物移走的影响,IVT算法跟踪误差明显增大. 随着遮挡物的出现和移走,DSSmap算法出现不同程度的跟踪漂移,其他采用局部稀疏表示模型的算法能有效处理遮挡干扰,基本实现了完整跟踪,但SCM算法和本文算法跟踪效果要优于其他算法.

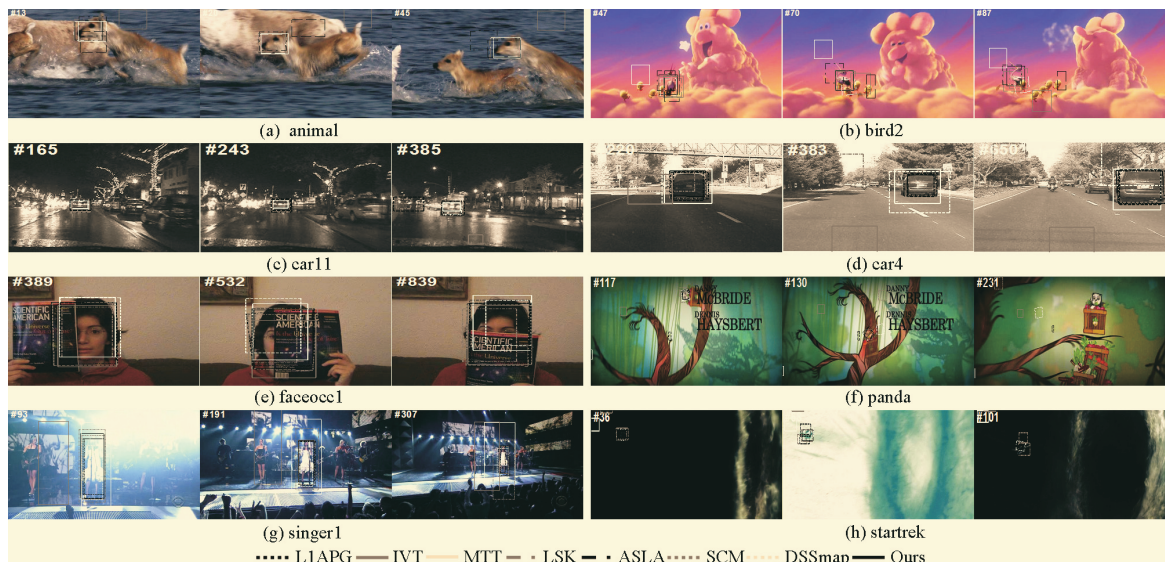


图2 部分时刻不同算法在每段测试视频序列上的跟踪效果

panda 测试序列受尺度变化和遮挡干扰等影响. 在第 10 帧目标出现运动模糊, LSK 算法跟踪误差明显增大, 本文算法在两帧后又重新实现准确跟踪, MTT 算法出现跟踪飘移. 在第 128、210 帧目标隐藏, SCM 算法和本文算法对目标的运动趋势跟踪基本正确. 在第 225 帧后, 目标尺寸增加且存在运动模糊, 本文算法跟踪性能明显优于 SCM 算法.

singer1 测试序列存在尺度变化和背景干扰等. LSK 算法简单利用跟踪结果更新模板容易引起跟踪飘移. SCM 算法和本文算法跟踪性能良好, SCM 算法在全局特征模板基础上运用判决模型从背景中区分出目标, 本文算法运用优化的稀疏字典可以更简单有

效地将目标从背景中分离出来.

startrek 测试序列存在剧烈运动变化、自身形变和严重的背景干扰, 目标一度淹没在急剧变化的背景中, 如在第 7、19、87 帧, LIAPG 算法跟踪误差明显增大, 本文算法一直都能准确跟踪目标. 虽然在第 108 帧本文算法出现了跟踪偏移, 但是能快速找准目标中心, 较其他算法有明显优势.

采用中心误差评价准则进行定量分析, 表 1 为不同算法在每段测试序列中的中心位置误差平均值, 图 3 为不同时刻不同算法在每段测试序列中的中心位置误差值.

表 1 中心位置误差平均值 (pixels)

算法	animal	bird2	car11	car4	faceoccl	panda	singer1	startrek
LIAPG	85.2869	60.7606	18.5365	4.1414	14.5673	93.0738	5.7793	15.6222
IVT	233.5120	49.0733	15.3923	98.3736	18.1992	131.3416	92.4879	100.3876
MTT	19.9091	102.1745	2.1754	7.2282	26.2420	257.7571	14.7661	90.1640
LSK	13.7866	16.8368	21.9423	114.4594	25.7146	65.5889	11.9826	8.5040
ASLA	26.0987	27.3592	2.2587	3.3071	15.5580	59.8093	4.6506	12.1147
SCM	10.7353	9.7074	2.1133	3.2666	12.5420	4.3163	2.2858	9.6892
DSSmap	9.4554	18.6778	2.1018	28.4739	27.5883	65.8142	11.3261	11.4054
Ours	12.0439	8.8792	2.1708	2.9799	12.3671	4.2853	3.9471	7.4087

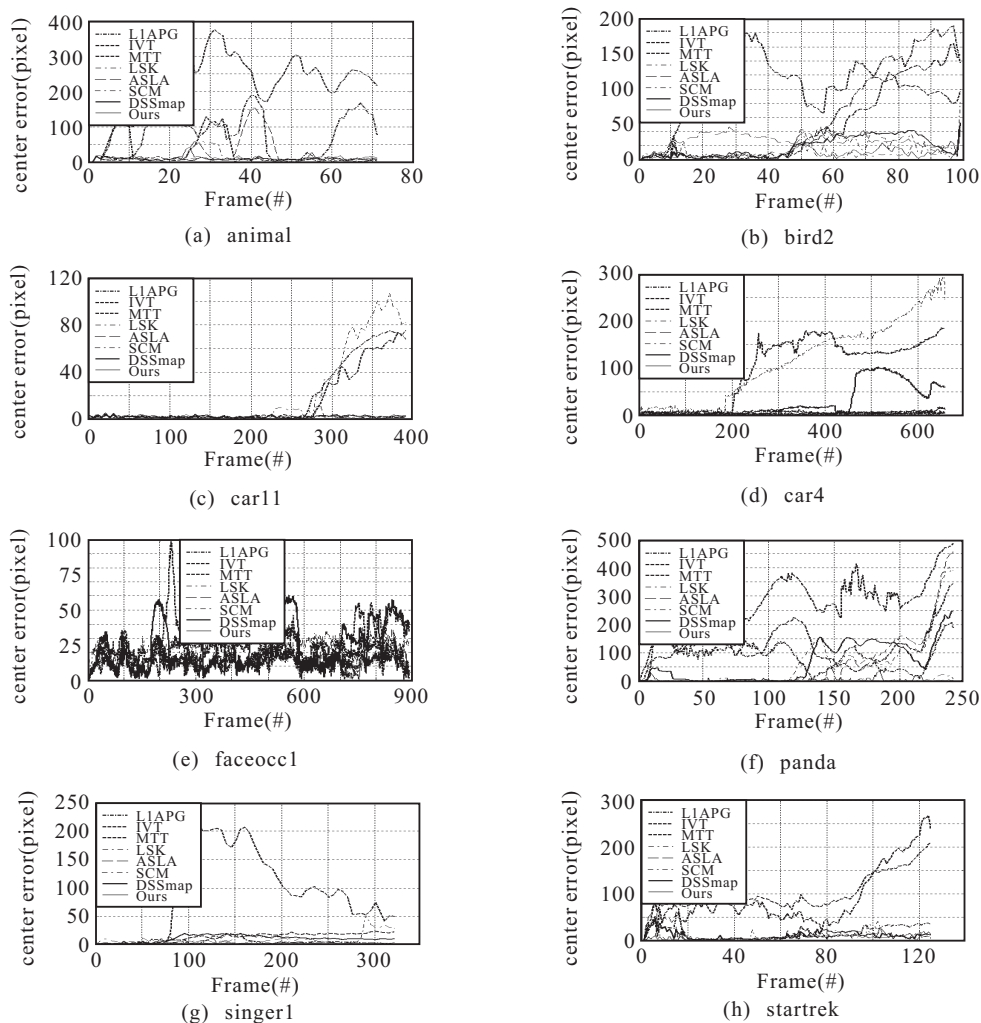


图 3 不同时刻不同算法在每段测试序列中的中心位置误差值

表 2 每帧平均耗时

算法	LIAPG	IVT	MTT	LSK	ASLA	SCM	DSSmap	Ours
每帧平均耗时	0.539 2	0.032 9	1.037 2	0.190 7	0.117 6	1.871 3	0.269 4	0.912 5

由表 1 可见, 本文算法在典型测试环境下的中心位置平均误差较小, 平均跟踪误差能控制在 13 pixels 以内, 跟踪性能明显优于 DSSmap 算法. 由图 2 和图 3 可见, 本文算法运用优化的稀疏字典能更加鲁棒地表征目标外观模型, 且空间信息相关性的充分利用保证了新算法能在自身形变和外界干扰等复杂情况下有效实现目标的稳定跟踪, 适用性好.

在算法运算耗时方面, 本文算法计算量与模板样本数目、候选样本数目成近似线性关系, 选择适当数目的候选样本可有效平衡消耗时间和性能提升, 本文算法中模板样本数目选为 10, 候选样本数目选为 200, 表 2 为各算法在测试序列中的每帧平均耗时.

由表 1 和表 2 可见, 本文算法中字典的优化学习与空间相关性信息的引入虽然在一定程度上牺牲了实时性, 但有效提高了算法跟踪性能.

5 结 论

本文结合空间结构信息和优化的稀疏字典, 设计了一种新的目标跟踪算法. 该算法利用似然估计理论和拉普拉斯正则项优化了目标代价函数中的稀疏性和空间结构相关性约束, 并借助拉格朗日对偶理论和加速近端梯度方法完成了稀疏字典优化; 然后, 基于最大池化理论和空间金字塔方法, 得到降维的且能包含更多空间信息的目标模板系数和候选样本系数; 最后, 利用相似性判定准则得到最佳候选样本, 实现目标准确跟踪. 实验结果表明, 所提出的算法能有效解决目标在具有背景干扰、光照变化、形变、运动模糊、严重遮挡等多种复杂场景中跟踪精度低的问题.

参考文献(References)

- [1] Wu Y, Lim J, Yang M H. Online object tracking: A benchmark[C]. Proc of IEEE Conf on Computer Vision and Pattern Recognition. Portland: IEEE Press, 2013: 2411-2418.
- [2] Kristan M, Pflugfelder R, Leonardis A, et al. The visual object tracking vot2014 challenge results[C]. Proc of the 13th European Conf on Computer Vision: Part II. Switzerland: Springer International Publishing, 2014: 191-217.
- [3] Mei X, Ling H. Robust visual tracking using l_1 minimization[C]. Proc of IEEE 12th Int Conf on Computer Vision. Kyoto: IEEE Press, 2009: 1436-1443.
- [4] Bao C, Wu Y, Ling H, et al. Real time robust l_1 tracker using accelerated proximal gradient approach[C]. Proc of IEEE Conf on Computer Vision and Pattern Recognition. Providence: IEEE Press, 2012: 1830-1837.

- [5] Jia X, Lu H, Yang M H. Visual tracking via adaptive structural local sparse appearance model[C]. Proc of IEEE Conf on Computer Vision and Pattern Recognition. Providence: IEEE Press, 2012: 1822-1829.
- [6] 张旭东, 陈仲海, 胡良梅, 等. 基于联合模板稀疏表示的目标跟踪方法[J]. 控制与决策, 2015, 30(7): 1696-1700. (Zhang X D, Chen Z H, Hu L M, et al. Object tracking method based on sparse representation of joint template[J]. Control and Decision, 2015, 30(7): 1696-1700.)
- [7] Zhong W, Lu H, Yang M H. Robust object tracking via sparsity-based collaborative model[C]. Proc of IEEE Conf on Computer Vision and Pattern Recognition. Providence: IEEE Press, 2012: 1838-1845.
- [8] Zhuang B, Lu H, Xiao Z, et al. Visual tracking via discriminative sparse similarity map[J]. IEEE Trans on Image Processing, 2014, 23(4): 1872-1881.
- [9] Yang M, Zhang L, Yang J, et al. Robust sparse coding for face recognition[C]. Proc of IEEE Conf on Computer Vision and Pattern Recognition. Providence: IEEE Press, 2011: 625-632.
- [10] Gao S, Tsang I W H, Chia L T, et al. Local features are not lonely—Laplacian sparse coding for image classification[C]. Proc of IEEE Conf on Computer Vision and Pattern Recognition. San Francisco: IEEE Press, 2010: 3555-3561.
- [11] Lee H, Battle A, Raina R, et al. Efficient sparse coding algorithms[C]. Advances in Neural Information Processing Systems. Vancouver: MIT Press, 2006: 801-808.
- [12] Yang J, Yu K, Gong Y, et al. Linear spatial pyramid matching using sparse coding for image classification[C]. Proc of IEEE Conf on Computer Vision and Pattern Recognition. Miami: IEEE Press, 2009: 1794-1801.
- [13] Ross D A, Lim J, Lin R S, et al. Incremental learning for robust visual tracking[J]. Int J of Computer Vision, 2008, 77(1/2/3): 125-141.
- [14] Zhang T, Ghanem B, Liu S, et al. Robust visual tracking via multi-task sparse learning[C]. Proc of IEEE Conf on Computer Vision and Pattern Recognition. Providence: IEEE Press, 2012: 2042-2049.
- [15] Liu B, Huang J, Kulikowski C, et al. Robust visual tracking using local sparse appearance model and k-selection[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2013, 35(12): 2968-2981.