

基于双向稀疏表示的鲁棒目标跟踪算法

王保宪 赵保军 唐林波 王水根 吴京辉

Robust visual tracking algorithm based on bidirectional sparse representation

Wang Bao-Xian Zhao Bao-Jun Tang Lin-Bo Wang Shui-Gen Wu Jing-Hui

引用信息 Citation: Acta Physica Sinica 63, 234201 (2014) DOI: 10.7498/aps.63.234201

在线阅读 View online: <http://dx.doi.org/10.7498/aps.63.234201>

当期内容 View Table of Contents: <http://wulixb.iphy.ac.cn/CN/volumn/home.shtml>

您可能感兴趣的其他文章

Articles you may be interested in

[目标跟踪中目标模型更新问题的半监督学习算法研究](#)

高文, 汤洋, 朱明

2014, 63(21): 214204. 全文: [PDF](#) (7485KB)

[基于深度玻尔兹曼模型的红外与可见光图像融合](#)

冯鑫, 李川, 胡开群

2014, 63(18): 184202. 全文: [PDF](#) (1144KB)

[米散射激光雷达剖面数据三次样条垂直水平插值法](#)

陈浩, 华灯鑫, 张毅坤, 朱承炫

2014, 63(15): 154204. 全文: [PDF](#) (5413KB)

[复杂背景下目标检测的级联分类器算法研究](#)

高文, 汤洋, 朱明

2014, 63(9): 094204. 全文: [PDF](#) (6905KB)

[基于结构张量的变分多源图像融合](#)

赵文达, 赵建, 续志军

2013, 62(21): 214204. 全文: [PDF](#) (3948KB)

基于双向稀疏表示的鲁棒目标跟踪算法*

王保宪 赵保军[†] 唐林波 王水根 吴京辉

(北京理工大学信息与电子学院, 北京 100081)

(2014年6月17日收到; 2014年7月8日收到修改稿)

目前, 基于稀疏表示的目标跟踪通常在目标模板集上重构候选样本的正向模型或者在候选样本集上描述目标模板的反向模型. 两个模型的共同点是均需计算候选样本与模板集合之间的稀疏相关系数矩阵. 基于此, 建立了一个双向联合稀疏表示的跟踪模型, 该模型通过 L_2 范数约束正反向稀疏相关系数矩阵达到一致收敛. 与之前的单向稀疏表示模型相比, 双向稀疏表示跟踪模型在正反向联合求解框架下可以更加充分地挖掘所有候选样本与模板集之间的稀疏映射关系, 并将稀疏映射表上对正负模板区分度最好的候选样本作为目标. 基于加速逼近梯度 (accelerated proximal gradient) 快速算法, 以矩阵形式推导了双向稀疏表示模型的求解框架, 使得候选样本集和目标模板集均以矩阵方式并行求解, 在一定程度上提高了计算效率. 实验数据表明所提出的算法优于传统的单向稀疏表示目标跟踪算法.

关键词: 稀疏表示, 目标跟踪, 贝叶斯框架

PACS: 42.30.Va, 07.05.Pj, 42.30.Tz, 02.50.-r

DOI: 10.7498/aps.63.234201

1 引言

目标跟踪一直以来是计算机视觉和图像处理领域的研究热点之一, 在视频监控、车辆导航、自动目标检测以及人机智能交互等方面得到广泛地应用^[1]. 尽管在目标跟踪领域已经有不少跟踪算法在一定程度上取得了良好的效果, 但仍有很多问题没有解决. 比如目标运动突变、目标遮挡、背景光照突变以及跟踪场景杂乱等因素均给目标的鲁棒跟踪带来很大的挑战.

最近, 稀疏表示理论^[2]已经成功应用到了人脸识别^[3]、图像重建^[4]以及图像分割^[5]等领域. Mei和Ling^[6]首次将稀疏表示引入到目标跟踪领域, 他们认为跟踪目标可以在目标模板和小模板构成的字典中线性稀疏表示, 其中, 目标模板用于重建跟踪目标, 而小模板用于描述背景随机噪声. 由此将目标跟踪问题看作为一个在贝叶斯推理框架下的 L_1 优化问题 (L_1 跟踪), 并通过计算候选样本在

目标模板集上的重构误差值估计目标. 其中 L_1 跟踪求解凸优化问题的次数和跟踪算法的收敛性均与样本个数成正相关, 所以计算量限制了 L_1 跟踪算法. 其后, Bao等^[7]应用一种快速梯度算法对 L_1 跟踪框架进行了一定程度的加速.

当目标外观发生突变时, 真实样本在目标模板集上的重构误差比较大, 此时 L_1 跟踪算法难以估计出最佳样本. 文献^[8]用跟踪目标附近的背景模板替换 L_1 跟踪字典中的小模板并将在目标模板和背景模板上对比差异最高的样本作为目标. 基于此, 文献^[8]提出了两个独立稀疏表示模型并综合两个模型的信息进行目标跟踪, 但是该算法需要对所有候选样本进行两次稀疏求解, 从而增加了跟踪算法的计算量.

以上跟踪算法均是将目标模板作为过完备字典的组成部分, 并通过添加稀疏约束条件挑选与候选样本线性相关的目标模板描述候选样本. 本文称这类模型为正向稀疏表示模型. 与正向稀疏表示模型相反的是将候选样本作为字典的组成部分, 并通

* 国家高技术研究发展计划 (批准号: 2012AA8012011C) 资助的课题.

[†] 通讯作者. E-mail: zbj@bit.edu.cn

过求解目标模板在候选样本集上的稀疏映射关系估计目标. 例如, Liu 和 Sun^[9] 在固定目标模板的基础上, 通过求解目标模板在所有候选样本和小模板构成的字典上的稀疏表示系数, 并将稀疏表示系数作为权重确定跟踪目标. 显然 Liu 的方法仅需求解 1 次 L_1 优化问题, 计算量大幅减少, 但是其固定的目标模板难以适应复杂的跟踪场景. 在文献 [8] 的基础上, 文献 [10] 反过来求解目标模板和背景模板在候选样本集上的稀疏表示系数, 然后, 结合模板和样本的距离关系建立一个概率相似图并由此确定目标样本. 但是该算法需要足量的候选样本以保证跟踪字典的完备性, 并且当目标跟踪出现偏差时, 更新的背景模板中会包含部分类似目标的模板, 导致概率相似图的可区分能力下降, 进而易发生模板漂移.

通过研究基于稀疏表示的目标跟踪框架, 发现正向稀疏表示模型是用目标模板的线性组合稀疏重构候选样本, 而反向稀疏表示模型是用候选样本的线性组合近似表达目标模板, 两者均需计算模板与样本之间的稀疏相关系数矩阵. 理论上, 模板与样本之间的相关度量是一个固有属性, 正向稀疏表示模型和反向稀疏表示模型得到的结果应该是一致的. 基于此, 本文将两者统一并建立了一种广义的双向稀疏表示跟踪模型. 该模型利用 L_2 范数约束正反向稀疏相关系数矩阵的差值, 从而使两者内容达到一个很好的收敛. 在正反向联合稀疏求解框架下, 双向稀疏表示模型能够充分地挖掘所有候选样本与模板之间的相关度量关系, 并将稀疏映射表上对正负模板区分度最好的候选样本作为目标. 此外为了提高算法执行效率, 本文采用了加速逼近梯度 (accelerated proximal gradient, APG) 快速解法并以矩阵形式推导了双向稀疏表示模型的求解框架, 从而使算法不像传统 L_1 跟踪算法那样以单个候选样本或模板进行稀疏求解, 在一定程度上提高了算法的计算效率. 实验数据表明, 双向稀疏表示模型比传统的单向稀疏表示模型具有更好的鲁棒性.

2 相关内容

2.1 贝叶斯推理估计

目标跟踪问题可以视为含隐藏状态变量的马尔可夫模型下的贝叶斯推理问题. 本文选

用 6 个仿射运动变换参数描述目标状态, 即有 $\mathbf{x} = \{r, c, \theta, s, \alpha, \phi\}$. 其中, r, c 分别表示目标在水平和垂直方向上的平移; θ, s, α, ϕ 分别表示目标的旋转角度、尺度变化值、比例变化值以及倾斜角度. 在给定 l 帧连续图像序列观测值 $\mathbf{Z}_l = \{z_1, z_2, \dots, z_l\}$ 后, 可按下式估计第 l 帧图像目标的状态值 \mathbf{x}_l :

$$p\left(\frac{\mathbf{x}_l}{\mathbf{Z}_l}\right) \propto p\left(\frac{z_l}{\mathbf{x}_l}\right) \int p\left(\frac{\mathbf{x}_l}{\mathbf{x}_{l-1}}\right) p\left(\frac{\mathbf{x}_{l-1}}{\mathbf{Z}_{l-1}}\right) d\mathbf{x}_{l-1}. \quad (1)$$

这里 $p(\mathbf{x}_l/\mathbf{x}_{l-1})$ 表示前后两帧目标状态的转移模型, 一般假设其服从高斯分布, 即有 $p(\mathbf{x}_l/\mathbf{x}_{l-1}) = N(\mathbf{x}_l; \mathbf{x}_{l-1}, \mathbf{\Omega})$, 其中 $\mathbf{\Omega}$ 为对角矩阵. $\mathbf{\Omega}$ 对角元素依次是 6 个仿射运动变换参数的方差, 即

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

在高斯状态分布下, 可以得到第 l 帧 n 个候选样本, 即

$$\mathbf{Y} = \{\mathbf{y}_i \in \mathbb{R}^{d \times 1}\} \quad (i = 1, 2, \dots, n), \quad (3)$$

这里 d 代表候选样本图像的特征维数. $p(z_l/\mathbf{x}_l)$ 表示观测似然模型, 用于估计图像 z_l 中目标状态为 \mathbf{x}_l 的概率, 本文中 $p(z_l/\mathbf{x}_l)$ 正比于候选样本在稀疏映射表上对正负模板的区分度. 可通过最大后验概率方法估计出目标的最佳状态值 $\hat{\mathbf{x}}_l$, 即

$$\hat{\mathbf{x}}_l = \arg \max_{\mathbf{x}_l^i} p(\mathbf{x}_l^i/\mathbf{Z}_l), \quad (4)$$

式中 \mathbf{x}_l^i 表示图像目标状态 \mathbf{x}_l 的第 i 个采样.

2.2 单向稀疏表示模型

设模板集合 $\mathbf{T} = \{\mathbf{t}_i \in \mathbb{R}^{d \times 1}\}, i = 1, 2, \dots, m$, 这里 d 表示模板图像的特征维数. 本文采用文献 [8] 的方法, 认为模板集合 \mathbf{T} 由 p 个目标模板 (正模板) \mathbf{T}_{pos} 和 q 个背景模板 (负模板) \mathbf{T}_{neg} 组成. 其中, 背景模板作为负模板描述非目标部分, 即

$$\begin{aligned} \mathbf{T} &= [\mathbf{T}_{\text{pos}}, \mathbf{T}_{\text{neg}}] \\ &= [\mathbf{t}_1, \dots, \mathbf{t}_p, \mathbf{t}_{p+1}, \dots, \mathbf{t}_{p+q}]. \end{aligned} \quad (5)$$

对于特征描述模型, 常见有颜色直方图、灰度、梯度方向直方图等^[11]. 本文采用灰度特征进行跟踪, 其他特征同样适用于本文算法.

正向稀疏表示模型认为在跟踪进程中, 候选样本 \mathbf{y} 可由模板集合 \mathbf{T} 近似线性表示且仅有几个模板被用来频繁地表示候选样本. 由此可通过添加稀

疏约束条件挑选出这些与候选样本线性相关的模板去描述样本, 即求解如下优化表达式:

$$\min_{\mathbf{a}} \{ \|\mathbf{y} - \mathbf{T}\mathbf{a}\|_2^2 + \mu \|\mathbf{a}\|_1 \}, \quad (6)$$

这里, $\|\cdot\|_2, \|\cdot\|_1$ 分别表示 $L2$ 范数和 $L1$ 范数; μ 是正则化参数量; $\mathbf{a} \in \mathbb{R}^{m \times 1}$, 是稀疏相关表示系数, 其代表了所有模板与候选样本 \mathbf{y} 的相关程度. 在求得 \mathbf{a} 后, 再通过相应的观测似然函数可计算出候选样本 \mathbf{y} 的接受概率.

反向稀疏表示模型认为, 在每帧重采样后, 候选样本集中仅有部分样本比较贴近目标状态, 即目标模板在候选样本集上呈现低秩映射. 因此可将候选样本集合 \mathbf{Y} 作为过完备字典, 求解模板 \mathbf{t} 在过完备字典上的稀疏表示, 即

$$\min_{\mathbf{b}} \{ \|\mathbf{t} - \mathbf{Y}\mathbf{b}\|_2^2 + \eta \|\mathbf{b}\|_1 \}, \quad (7)$$

这里, η 同样是正则化参数量; $\mathbf{b} \in \mathbb{R}^{n \times 1}$, 是模板 \mathbf{t} 在候选样本集上的稀疏相关表示系数, 代表了该模板与所有候选样本的相关程度. 当 \mathbf{t} 为目标模板时, 可以将稀疏相关表示系数近似为候选样本的接受概率.

3 双向稀疏表示目标跟踪

3.1 模型建立

在正向稀疏表示模型下, 可依次求解候选样本 \mathbf{y}_i 在模板集合 \mathbf{T} 上的稀疏表示:

$$\begin{cases} \min_{\mathbf{a}_1} \{ \|\mathbf{y}_1 - \mathbf{T}\mathbf{a}_1\|_2^2 + \mu \|\mathbf{a}_1\|_1 \}, \\ \vdots \\ \min_{\mathbf{a}_{n-1}} \{ \|\mathbf{y}_{n-1} - \mathbf{T}\mathbf{a}_{n-1}\|_2^2 + \mu \|\mathbf{a}_{n-1}\|_1 \}, \\ \min_{\mathbf{a}_n} \{ \|\mathbf{y}_n - \mathbf{T}\mathbf{a}_n\|_2^2 + \mu \|\mathbf{a}_n\|_1 \}. \end{cases}$$

为了提高计算效率, 可将所有候选样本整体进行求解, 即

$$\min_{\mathbf{A}} \{ \|\mathbf{Y} - \mathbf{T}\mathbf{A}\|_2^2 + \mu \|\mathbf{A}\|_1 \}, \quad (8)$$

这里 $\mathbf{A} \in \mathbb{R}^{m \times n}$, 为候选样本集 \mathbf{Y} 在模板集上的稀疏相关系数矩阵, 其代表了所有模板集 \mathbf{T} 与候选样本集 \mathbf{Y} 的相关程度. 矩阵 \mathbf{A} 中的元素均是非负的, 即有 $\mathbf{A} \geq \mathbf{0}$. 同样在反向稀疏表示模型下, 也可整体求解所有模板在候选样本集上的稀疏表示,

$$\min_{\mathbf{B}} \{ \|\mathbf{T} - \mathbf{Y}\mathbf{B}\|_2^2 + \eta \|\mathbf{B}\|_1 \}, \quad (9)$$

这里 $\mathbf{B} \in \mathbb{R}^{n \times m}$, 为模板集 \mathbf{T} 在候选样本集上的稀疏相关系数矩阵, 表征了所有候选样本集 \mathbf{Y} 与模板集 \mathbf{T} 的相关度量. 矩阵 \mathbf{B} 中的所有元素也是非负的, 即 $\mathbf{B} \geq \mathbf{0}$.

以上是传统正向稀疏表示和反向稀疏表示模型的矩阵表达式. 显然, 两者的共同点是均需计算候选样本集 \mathbf{Y} 和模板集 \mathbf{T} 的稀疏相关系数矩阵. 虽然正反向稀疏表示模型均取得了较好的跟踪效果, 但其仍存在以下两个问题. 首先, 稀疏相关系数矩阵是通过几次迭代收敛得到的, 其与真实的最优解仍存在一定的误差, 故得到的跟踪结果并不是

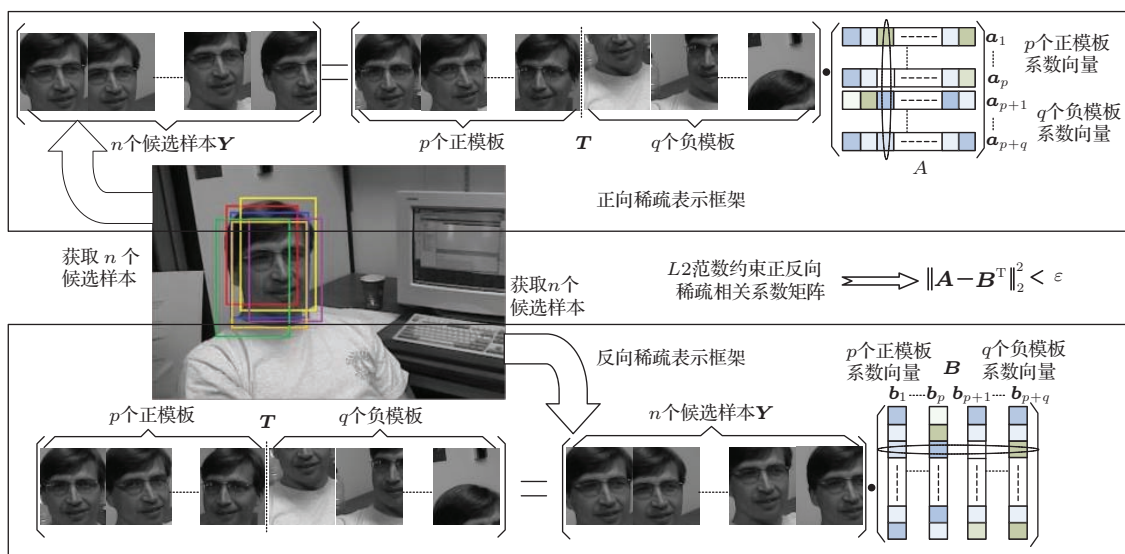


图1 (网刊彩色) 双向稀疏表示模型的建立过程

最佳的. 其次, 候选样本 \mathbf{Y} 与模板集合 \mathbf{T} 之间的相关性是一个固有属性, 所以稀疏相关系数矩阵 \mathbf{A} 和 \mathbf{B} 的内容应该是一致的. 但在单向稀疏表示框架下两者是分别独立求解的, 无法考虑 \mathbf{A} 和 \mathbf{B} 内在的关联性, 因而导致跟踪算法无法将两者结合提高跟踪算法的稳定性. 为了解决这些问题, 本文在正反向稀疏相关系数矩阵的求解基础上构造了一个损失函数来约束正反向稀疏求解, 即有如下优化框架:

$$\begin{cases} \min_{\mathbf{A}} \left\{ \|\mathbf{Y} - \mathbf{T}\mathbf{A}\|_2^2 + \mu \|\mathbf{A}\|_1 \right\}, \\ \min_{\mathbf{B}} \left\{ \|\mathbf{T} - \mathbf{Y}\mathbf{B}\|_2^2 + \eta \|\mathbf{B}\|_1 \right\}, \\ \text{s.t. } \|\mathbf{A} - \mathbf{B}^T\|_2^2 < \varepsilon. \end{cases} \quad (10)$$

这里 ε 是个比较小的常数参量, 用于描述正反向稀疏相关系数矩阵的差值. 损失函数利用 L_2 范数从能量角度约束正向稀疏表示和反向稀疏表示的最优解达到一致收敛. 图 1 展示了双向稀疏表示模型建立的具体过程.

为了使正反向稀疏表示建立在一个优化求解函数下, 本文采用拉格朗日因子法将损失函数作为新的约束项, 从而建立一个双向联合稀疏表示的目标跟踪框架:

$$\begin{aligned} & \min_{\mathbf{A}, \mathbf{B}} \left\{ \|\mathbf{Y} - \mathbf{T}\mathbf{A}\|_2^2 + \|\mathbf{T} - \mathbf{Y}\mathbf{B}\|_2^2 \right. \\ & \left. + \lambda \|\mathbf{A} - \mathbf{B}^T\|_2^2 + \mu \|\mathbf{A}\|_1 + \eta \|\mathbf{B}\|_1 \right\}, \\ & \text{s.t. } \mathbf{A} \geq \mathbf{0}, \mathbf{B} \geq \mathbf{0}, \end{aligned} \quad (11)$$

其中 λ 是正则化参数量. 由于双向联合稀疏求解框架充分考虑了 \mathbf{A} 和 \mathbf{B} 内在的关联性, 因而在迭代收敛中不仅能确保候选样本在模板集中具有良好的稀疏表达, 同时能够促使模板集合在候选样本集上达到低秩映射.

3.2 优化求解

本文基于 APG 快速算法求解双向联合稀疏表示优化问题. APG 算法^[12] 起初是用来求解如下无约束优化问题:

$$\omega = \arg \min_{\omega} F(\omega) = \arg \min_{\omega} \{f(\omega) + g(\omega)\}, \quad (12)$$

这里, $f(\omega)$ 是光滑可微的凸函数且有 Lipschitz 常数, $g(\omega)$ 是不光滑的凸函数项. 对于这种函数组合的优化求解, 应用 APG 算法可以达到二次收敛的速度.

3.2.1 APG 算法基本框架

算法 1 APG 算法框架

步骤 1 设最优解初始估计值: $\omega_0 = \omega_{-1} = 0$, 迭代因子 $\xi_0 = \xi_{-1}$.

步骤 2 依次迭代求解下面的表达式, $k = 0, 1, \dots$, 直至收敛.

$$\begin{cases} \beta_{k+1} = \omega_k + \frac{\xi_{k-1} - 1}{\xi_k} (\omega_k - \omega_{k-1}), \\ \omega_{k+1} = \arg \min_{\omega} \left\{ \frac{L}{2} \left\| \omega - \beta_{k+1} + \frac{\nabla f(\beta_{k+1})}{L} \right\|_2^2 \right. \\ \left. + g(\omega) \right\}, \\ \xi_{k+1} = \frac{1 + \sqrt{1 + 4\xi_k^2}}{2}. \end{cases}$$

这里, β_{k+1} 是第 k 次迭代估计的中间回归量, ξ_k 是第 k 次迭代的加速收敛因子, L 是函数 $f(\omega)$ 的 Lipschitz 常数. 由于优化问题 (11) 式带有非负约束条件, 故其无法直接应用 APG 算法进行求解. 本文通过一种指示函数将其转换为无条件约束的优化问题. 指示函数 $\phi(\omega)$ 的定义如下:

$$\phi(\omega) = \begin{cases} 0 & (\omega \geq 0), \\ +\infty & (\omega < 0). \end{cases}$$

应用指示函数, 可将优化问题 (11) 式等价为

$$\begin{aligned} & \min_{\mathbf{A}, \mathbf{B}} \left\{ \|\mathbf{Y} - \mathbf{T}\mathbf{A}\|_2^2 + \|\mathbf{T} - \mathbf{Y}\mathbf{B}\|_2^2 + \lambda \|\mathbf{A} - \mathbf{B}^T\|_2^2 \right. \\ & \left. + \mu \|\mathbf{A}\|_1 + \eta \|\mathbf{B}\|_1 + \phi(\mathbf{A}) + \phi(\mathbf{B}) \right\}. \end{aligned} \quad (13)$$

APG 算法的重点是求解步骤 2 中的优化子问题. 在引入指示函数后, 优化子问题具有如下解析解:

$$\omega_{k+1} = \max \left(\mathbf{0}, \beta_{k+1} + \frac{\nabla f(\beta_{k+1})}{L} \right). \quad (14)$$

此时, 应用 APG 算法可以在几次迭代中快速求得最优解. 由于优化问题 (13) 式具有两个未知变量, 因而没有确定的闭合解析解. 本文设计了一个交替更新 \mathbf{A} 和 \mathbf{B} 的迭代优化框架. 定义 $\mathbf{1}_Y \in \mathbb{R}^{n \times 1}$ 和 $\mathbf{1}_T \in \mathbb{R}^{m \times 1}$, 均为元素值均为 1 的列向量.

3.2.2 固定 \mathbf{A} , 更新 \mathbf{B}

在固定正向稀疏相关系数矩阵 \mathbf{A} 时, 更新求解 \mathbf{B} . 此时优化问题 (13) 式等价于

$$\min_{\mathbf{B}} \left\{ \|\mathbf{T} - \mathbf{Y}\mathbf{B}\|_2^2 + \lambda \|\mathbf{A} - \mathbf{B}^T\|_2^2 + \eta \|\mathbf{B}\|_1 + \phi(\mathbf{B}) \right\}. \quad (15)$$

在 APG 求解框架下, 将 (15) 式进行分解得到

$$f(\mathbf{B}) = \|\mathbf{T} - \mathbf{Y}\mathbf{B}\|_2^2 + \lambda \|\mathbf{A} - \mathbf{B}^T\|_2^2 + \eta \|\mathbf{B}\|_1,$$

$$g(\mathbf{B}) = \phi(\mathbf{B}).$$

基于 APG 算法框架, 可以推导出求解反向稀疏相关系数矩阵 \mathbf{B} 的计算步骤.

算法 2 求解反向稀疏相关系数矩阵 \mathbf{B}

步骤 1 设最优解初始估计值: $\mathbf{B}^0 = \mathbf{B}^{-1} = \mathbf{B}_\Delta$, 迭代因子 $\xi_0 = \xi_{-1} = 1$.

步骤 2 依次迭代求解下面这个表达式, $k = 0, 1, \dots$, 直至收敛.

$$\left\{ \begin{array}{l} \mathbf{V}_B^{(k)} = \mathbf{B}^{(k)} + \frac{\xi_{k-1} - 1}{\xi_k} (\mathbf{B}^{(k)} - \mathbf{B}^{(k-1)}), \\ \mathbf{H}_B = \mathbf{V}_B^{(k)} - \frac{1}{L_B} \left(2\mathbf{Y}^T (\mathbf{Y} \mathbf{V}_B^{(k)} - \mathbf{T}) \right. \\ \quad \left. + 2\lambda (\mathbf{V}_B^{(k)} - \mathbf{A}^T) + \eta \mathbf{1}_Y \mathbf{1}_T^T \right), \\ \mathbf{B}^{(k+1)} = \max(\mathbf{0}, \mathbf{H}_B), \\ \xi_{k+1} = \frac{1 + \sqrt{1 + 4\xi_k^2}}{2}. \end{array} \right.$$

这里, \mathbf{B}_Δ 是整个框架对 \mathbf{B} 交替迭代更新的上次估计值, $\mathbf{V}_B^{(k)}$ 是步骤 2 中 APG 内层循环对 \mathbf{B} 进行第 k 次迭代的中间回归量, \mathbf{H}_B 是步骤 2 中 APG 内层循环对 \mathbf{B} 优化子问题的解析解, L_B 是函数 $f(\mathbf{B})$ 的 Lipschitz 常数.

3.2.3 固定 \mathbf{B} , 更新 \mathbf{A}

在固定反向稀疏相关系数矩阵 \mathbf{B} 时, 更新求解 \mathbf{A} . 此时优化问题 (13) 式等价于

$$\min_{\mathbf{A}} \left\{ \|\mathbf{Y} - \mathbf{T}\mathbf{A}\|_2^2 + \lambda \|\mathbf{A} - \mathbf{B}^T\|_2^2 + \mu \mathbf{A} + \phi(\mathbf{A}) \right\}. \quad (16)$$

在 APG 求解框架下, 将 (16) 式进行分解得到

$$\begin{aligned} f(\mathbf{A}) &= \|\mathbf{Y} - \mathbf{T}\mathbf{A}\|_2^2 + \lambda \|\mathbf{A} - \mathbf{B}^T\|_2^2 + \mu \mathbf{A}, \\ g(\mathbf{A}) &= \phi(\mathbf{A}). \end{aligned}$$

类似于算法 2, 可以推出求解正向稀疏相关系数矩阵 \mathbf{A} 的计算步骤.

算法 3 求解正向稀疏相关系数矩阵 \mathbf{A}

步骤 1 设最优解初始估计值: $\mathbf{A}^0 = \mathbf{A}^{-1} = \mathbf{A}_\Delta$, 迭代因子 $\xi_0 = \xi_{-1} = 1$.

步骤 2 依次迭代求解下面这个表达式, $k = 0, 1, \dots$, 直至收敛.

$$\left\{ \begin{array}{l} \mathbf{V}_A^{(k)} = \mathbf{A}^{(k)} + \frac{\xi_{k-1} - 1}{\xi_k} (\mathbf{A}^{(k)} - \mathbf{A}^{(k-1)}), \\ \mathbf{H}_A = \mathbf{V}_A^{(k)} - \frac{1}{L_A} \left(2\mathbf{T}^T (\mathbf{T} \mathbf{V}_A^{(k)} - \mathbf{Y}) \right. \\ \quad \left. + 2\lambda (\mathbf{V}_A^{(k)} - \mathbf{B}^T) + \mu \mathbf{1}_T \mathbf{1}_Y^T \right), \\ \mathbf{A}^{(k+1)} = \max(\mathbf{0}, \mathbf{H}_A), \\ \xi_{k+1} = \frac{1 + \sqrt{1 + 4\xi_k^2}}{2}. \end{array} \right.$$

这里, \mathbf{A}_Δ 是整个框架对 \mathbf{A} 交替迭代更新的上次估计值, $\mathbf{V}_A^{(k)}$ 是步骤 2 中 APG 内层循环对 \mathbf{A} 进行第 k 次迭代的中间回归量, \mathbf{H}_A 是步骤 2 中 APG 内层循环对 \mathbf{A} 优化子问题的解析解, L_A 是函数 $f(\mathbf{A})$ 的 Lipschitz 常数.

理论上, 稀疏相关系数矩阵 \mathbf{A} 和 \mathbf{B} 交替更新框架需要迭代进行多次. 但实验发现相互迭代更新 1 次后, \mathbf{A} 和 \mathbf{B} 基本达到一致收敛.

3.3 观测似然模型

由 3.1 节可知, 稀疏相关系数矩阵描述了候选样本集 \mathbf{Y} 与模板集 \mathbf{T} 的相关程度. 而模板集 \mathbf{T} 由正负模板构成, 因此可将矩阵 \mathbf{A} 的每一列作为候选样本在正负模板上的区分特征量. 同样也可将矩阵 \mathbf{B} 的每一行作为该样本在正负模板上的区分特征量. 此外, 为了提高样本在正负模板上的区分能力, 本文对相关系数矩阵中的每个元素进行了距离权重计算. 设距离权重矩阵为 $\mathbf{W} \in \mathbb{R}^{m \times n}$, 定义如下:

$$W_{ij} \propto \exp \left(- \sum H_\delta(\mathbf{t}_i - \mathbf{y}_j) \right), \quad (17)$$

这里, $H_\delta(\cdot)$ 是 Huber 函数, 定义为

$$H_\delta(e) = \begin{cases} 0.5 \times e^2 & (|e| < \delta), \\ \delta|e| - 0.5 \times e^2 & (\text{其他}), \end{cases}$$

其中 δ 是 Huber 函数的分段调节参数. 显然当样本与模板比较相近时, 对应的权重值比较大; 反之对应的权重比较小. 与文献 [10] 采用的欧式距离函数相比, 本文采用的 Huber 函数具有更好的鲁棒性. 结合距离权重矩阵 \mathbf{W} , 本文将正反向稀疏相关系数矩阵进行并联组合, 得到 4 个不同的稀疏映射表:

$$\left\{ \begin{array}{l} \mathbf{M}^1 = \mathbf{A} \odot \mathbf{W}, \\ \mathbf{M}^2 = \mathbf{B}^T \odot \mathbf{W}, \\ \mathbf{M}^3 = \mathbf{A} \odot \mathbf{B}^T \odot \mathbf{W}, \\ \mathbf{M}^4 = (\mathbf{A} + \mathbf{B}^T) \odot \mathbf{W}. \end{array} \right.$$

这里 \odot 表示 Hadamard 乘积, 每个稀疏映射表均可表示为

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}_{1,\text{pos}} & \cdots & \mathbf{M}_{n,\text{pos}} \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ \mathbf{M}_{1,\text{neg}} & \cdots & \mathbf{M}_{n,\text{neg}} \end{bmatrix},$$

稀疏映射表中 $\mathbf{M}_{i,\text{pos}} \in \mathbb{R}^p$ 和 $\mathbf{M}_{i,\text{neg}} \in \mathbb{R}^q$, 分别表示候选样本 \mathbf{y}_i 在正负模板上的区分特征量. 可按下式计算候选样本 \mathbf{y}_i 在正负模板上的区分度:

$$d_i = \sum \mathbf{M}_{i,\text{pos}} - \sum \mathbf{M}_{i,\text{neg}}. \quad (18)$$

由 2.1 节可知, 在贝叶斯推理框架下将 d_i 作为该样本的观测似然值, 即

$$p(\mathbf{y}_l | \mathbf{x}_l) = \max p(\mathbf{y}_i | \mathbf{x}_l) = \max \{d_i\}. \quad (19)$$

根据最大后验概率可以确定出每个稀疏映射表下对应的目标样本, 由 4 个不同的稀疏映射表可以得到 4 个估计结果. 然后, 进一步在筛选出的 4 个估计样本中挑选与正模板集相似度最高的样本作为最终目标. 这种方法使得目标样本不仅在正负模板集上具有良好区分度, 而且与正模板集的距离达到最近, 因而求得的跟踪结果更为鲁棒.

3.4 模板更新

由 2.2 节可知, 模板集合由正模板 (目标模板) 和负模板 (背景模板) 组成. 本文采用文献 [10] 中的方法更新正模板内容. 对于负模板的更新, 当跟踪中出现遮挡等干扰时, 跟踪结果会出现一定程度的偏差. 此时持续更新的负模板中会包含部分目标信息, 易发生模板漂移. 为了解决这个问题, 本文通过目标跟踪结果反馈控制负模板的更新. 设第 l 帧跟踪结果在正负模板上的区分度为 d_l , 其前 5 帧跟踪结果在正负模板上的区分度均值为

$$\bar{d} = \frac{1}{5} (d_{l-1} + d_{l-2} + d_{l-3} + d_{l-4} + d_{l-5}). \quad (20)$$

定义第 l 帧跟踪结果在正负模板上的区分变化度为

$$\nu_l = \|d_l - \bar{d}\| / \bar{d}. \quad (21)$$

实验发现, 在稳定跟踪中跟踪结果在正负模板上的区分度是平稳变化的. 当跟踪出现干扰时, ν_l 值会出现突变. 如果 ν_l 大于阈值 ψ 时, 可判断当前帧出现跟踪干扰进而控制负模板不更新. 在后续 5 帧中, 控制 \bar{d} 值不变并当检查到新的跟踪结果使得 ν_l 小于阈值 ψ 时, 恢复负模板更新.

3.5 整体跟踪框架

在贝叶斯推理框架下, 本文利用 L_2 范数项约束正向稀疏相关系数矩阵 \mathbf{A} 和反向稀疏相关系数矩阵 \mathbf{B} 达到一致收敛, 并基于 APG 算法设计了一个交替更新 \mathbf{A} 和 \mathbf{B} 的迭代优化算法. 本文整体的跟踪框架如下.

算法 4 整体目标跟踪框架

- 1) 输入: N 帧跟踪视频, 首帧目标状态 \mathbf{x}_0 ;
- 2) 初始化: 获取模板集合 \mathbf{T} ; Lipschitz 常数 L_A, L_B ; 稀疏相关系数矩阵初始值 $\mathbf{A} = \mathbf{0}, \mathbf{B} = \mathbf{0}, \mathbf{A}_\Delta = \mathbf{0}, \mathbf{B}_\Delta = \mathbf{0}$; 视频帧号 $f = 1$;
- 3) 判断帧号 f 不大于 N , 开始处理第 f 帧图像; 否则, 跳至步骤 12;
- 4) 根据目标估计位置, 采集候选样本 \mathbf{Y} ;
- 5) 根据算法 3, 更新 \mathbf{A} , 并令 $\mathbf{A}_\Delta = \mathbf{A}$;
- 6) 根据算法 2, 更新 \mathbf{B} , 并令 $\mathbf{B}_\Delta = \mathbf{B}$;
- 7) 重复步骤 5: 根据算法 3, 更新得到 \mathbf{A} ;
- 8) 计算距离权重矩阵 \mathbf{W} ;
- 9) 根据观测似然模型, 估计出目标样本;
- 10) 对模板集进行更新;
- 11) 保留当前帧目标跟踪结果, 帧号 f 加 1, 返回步骤 3;
- 12) 输出整个视频的目标跟踪结果.

4 实验结果与分析

本文算法采用 Matlab 软件编写, 计算机平台配置为 Intel I7 (CPU3.0 GHz, 内存 16 GB). 本文选取 600 个候选样本, 10 个正模板以及 150 个负模板. 优化求解参数 λ 值为 0.08, μ 值和 η 值均为 0.04, Huber 函数的分段调节参数 δ 为 0.5, 模板更新判断阈值 ψ 为 0.7. 所有 APG 算法的内层循环为 5 次.

为了评估算法的性能, 我们与 4 种目前具有代表性的跟踪算法进行对比, 它们分别是: 加速逼近梯度 L_1 跟踪 (APGL1) [7], 压缩跟踪 (CT) [13], 区分概率图跟踪 (DSSP) [10] 和多示例学习跟踪 (MIL) [14]. 其中, APGL1 算法是基于正向稀疏表示的算法, 而 DSSP 是基于反向稀疏表示框架的算法. 为保证对比结果准确, 与本文对比的 4 种跟踪算法代码均在作者提供的网站上获得且所有跟踪算法在初始帧对跟踪目标的定位是相同的. 对比实

验在8组公开视频^[15]上进行,并从定性和定量两个方面对比衡量5种跟踪算法的跟踪效果.

4.1 定性分析

图2展示了5种跟踪算法在8个公开视频上的部分跟踪结果.在所有分图中,红色框代表本文算法的跟踪结果,而绿色框、蓝色框、黄色框和紫色框分别代表APGL1, CT, DSSP和MIL的跟踪结果.

图2(a)–(d)是boy视频和Deer视频的部分跟踪结果.两个视频均是跟踪快速运动的目标物,需要处理快速姿态变化、背景干扰、尺度变化等问题.仅有本文算法和DSSP算法成功完成了两个视频的跟踪任务.对于boy视频,由于APGL1算法不能

很好地处理目标快速运动变化,在第150帧时就跟丢目标.在boy后期视频中目标运动变化加剧,此时MIL算法和CT算法误差不断累积,在第523帧时发生模板漂移.在Deer视频第5帧中,跟踪目标出现快速运动,APGL1算法、CT算法和MIL算法均出现模板漂移.且由于目标运动变化加剧和背景信息干扰,三种跟踪算法的跟踪误差不断累积,在第38帧时彻底跟丢目标.

Car4视频是跟踪一个运动车辆,其包含了光照变化、背景干扰等复杂因素.图2(e)中当跟踪车辆经历光照变化时,CT算法和MIL算法的跟踪框不断变大.不断变大的跟踪框引入背景信息干扰,最终导致CT算法和MIL算法跟丢目标,如图2(f).

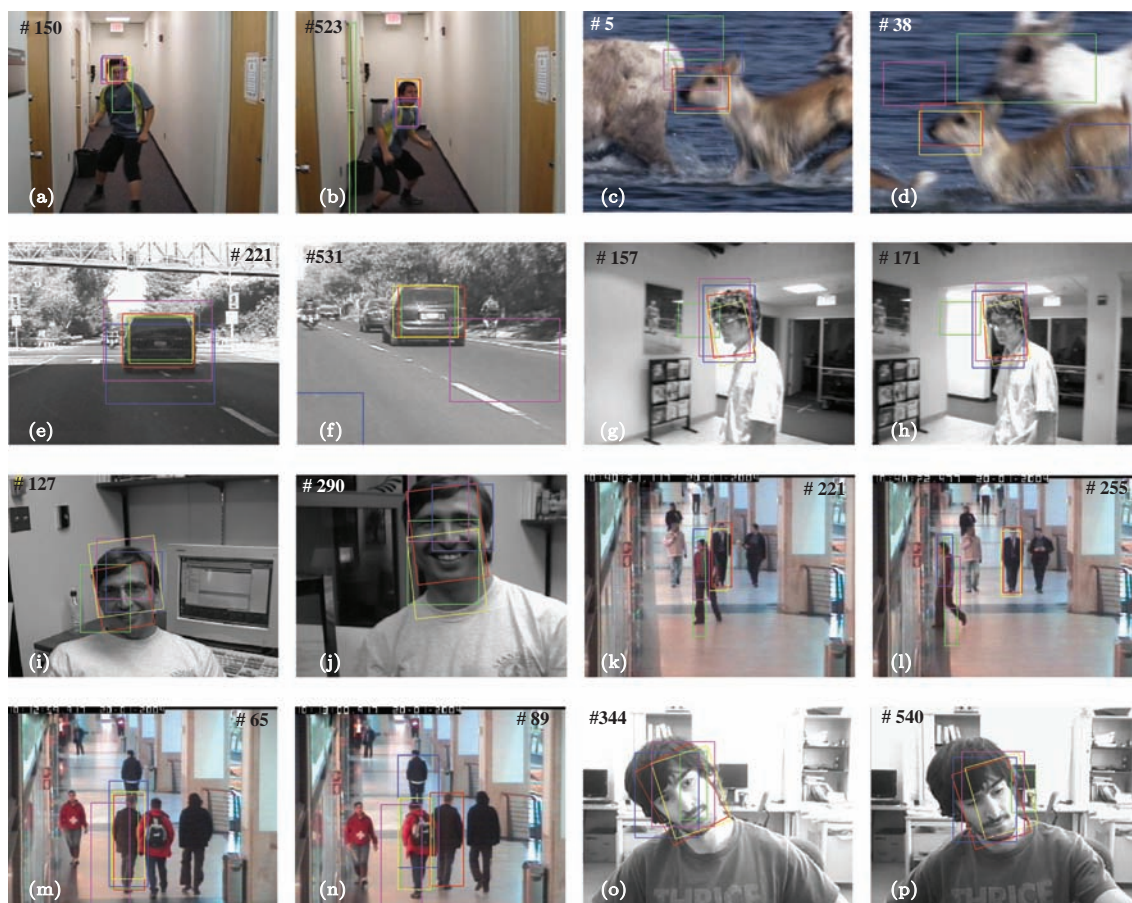


图2 (网刊彩色) 5种跟踪算法的部分跟踪结果 (a), (b)分别是boy第150, 523帧实验结果; (c), (d)分别是Deer第5, 38帧实验结果; (e), (f)分别是Car4第221, 531帧实验结果; (g), (h)分别是DavidIndoor第157, 171帧实验结果; (i), (j)分别是Dudek第127, 290帧实验结果; (k), (l)分别是Caviar2第221, 255帧实验结果; (m), (n)分别是Caviar3第65, 89帧实验结果; (o), (p)分别是Occlusion2第344, 540帧实验结果

DavidIndoor视频和Dudek视频均是跟踪室内运动的人脸,部分跟踪结果如图2(g)–(j).对于DavidIndoor视频,当人脸姿态出现快速变化时,APGL1算法不能及时调整导致跟踪框漂移.此外,

CT算法和MIL算法能完成大部分跟踪任务,但相比于本文算法和DSSP算法,其跟踪结果的中心误差值比较大.对于Dudek视频,仅有本文算法对视频进行了完整跟踪.在第127帧中,在跟踪目标出

现短时遮挡后,其他4种跟踪算法均出现了不同程度的模板漂移.在第290帧,当目标同时出现快速运动和尺度变化时,其他4种跟踪算法均出现了跟踪误差,而本文算法仍可以鲁棒地进行目标跟踪.

图2(k)—(p)展示了Caviar2, Caviar3以及Occlusion2视频的部分跟踪结果,三个视频均是为了验证跟踪算法处理目标遮挡的能力.在Caviar3视频中,跟踪目标发生了多次遮挡.在第89帧,其他4种跟踪算法均跟丢了目标,而本文算法仍继续跟踪初始目标.在Occlusion2视频中,跟踪目标在发生遮挡的同时出现了目标旋转,其他4种跟踪算法均出现了不同程度的跟踪偏差,仅有本文算法完成了精准跟踪.

4.2 定量分析

为了更加精准地对比跟踪算法的性能,本节首先计算了5种算法的跟踪结果与人工标记结果的中心误差值,并绘制了其随帧号变化的曲线图,具

体结果如图3所示.在所有分图中,本文算法的跟踪结果对应红色曲线,APGL1, CT, DSSP和MIL算法的跟踪结果分别对应绿色、蓝色、黄色和紫色曲线.

本文进一步计算了5种跟踪算法的平均中心定位误差值,如表1所列.平均中心定位误差值越小,则跟踪算法的定位精度越准.表1中,误差值最小的已经加粗表示.

下面给定人工标记的跟踪结果 R_G 和具体跟踪算法得到的跟踪结果 R_T ,可按照下式计算重叠率:

$$\text{overlap} = \frac{\text{area}(R_T \cap R_G)}{\text{area}(R_T \cup R_G)}. \quad (22)$$

一般重叠率越大,则表明跟踪结果与目标真实位置越接近.表2分别描述了5种跟踪算法的平均重叠率,平均重叠率最大的已经加粗表示.

从对比结果可以看出,本文算法在中心定位误差和重叠率两方面均优于其他4种跟踪算法,这说明了本文算法的跟踪性能好于其他算法.

表1 5种跟踪算法的平均中心定位误差(像素个数)

视频	APGL1	CT	DSSP	MIL	本文算法
boy	71.5	12.32	3.31	12.25	2.97
Car4	3.2	222.51	2.58	60.11	2.51
Caviar2	65.7	61.7	3.6	70.3	4.35
Caviar3	66.5	58.1	63.8	101.7	2.41
DavidIndoor	45.1	13.7	5.2	16.1	11
Deer	135.9	220.1	9.6	66.4	9.4
Dudek	84.2	23.5	112.3	35.7	19.2
Occlusion2	28.9	11.6	7	14.1	6.6

表2 5种跟踪算法的平均重叠率

视频	APGL1	CT	DSSP	MIL	本文算法
boy	0.19	0.54	0.77	0.51	0.79
Car4	0.87	0.21	0.9	0.34	0.92
Caviar2	0.33	0.31	0.75	0.26	0.76
Caviar3	0.21	0.18	0.19	0.06	0.86
DavidIndoor	0.23	0.49	0.67	0.45	0.61
Deer	0.08	0.04	0.61	0.2	0.62
Dudek	0.29	0.51	0.35	0.48	0.55
Occlusion2	0.31	0.63	0.72	0.61	0.78

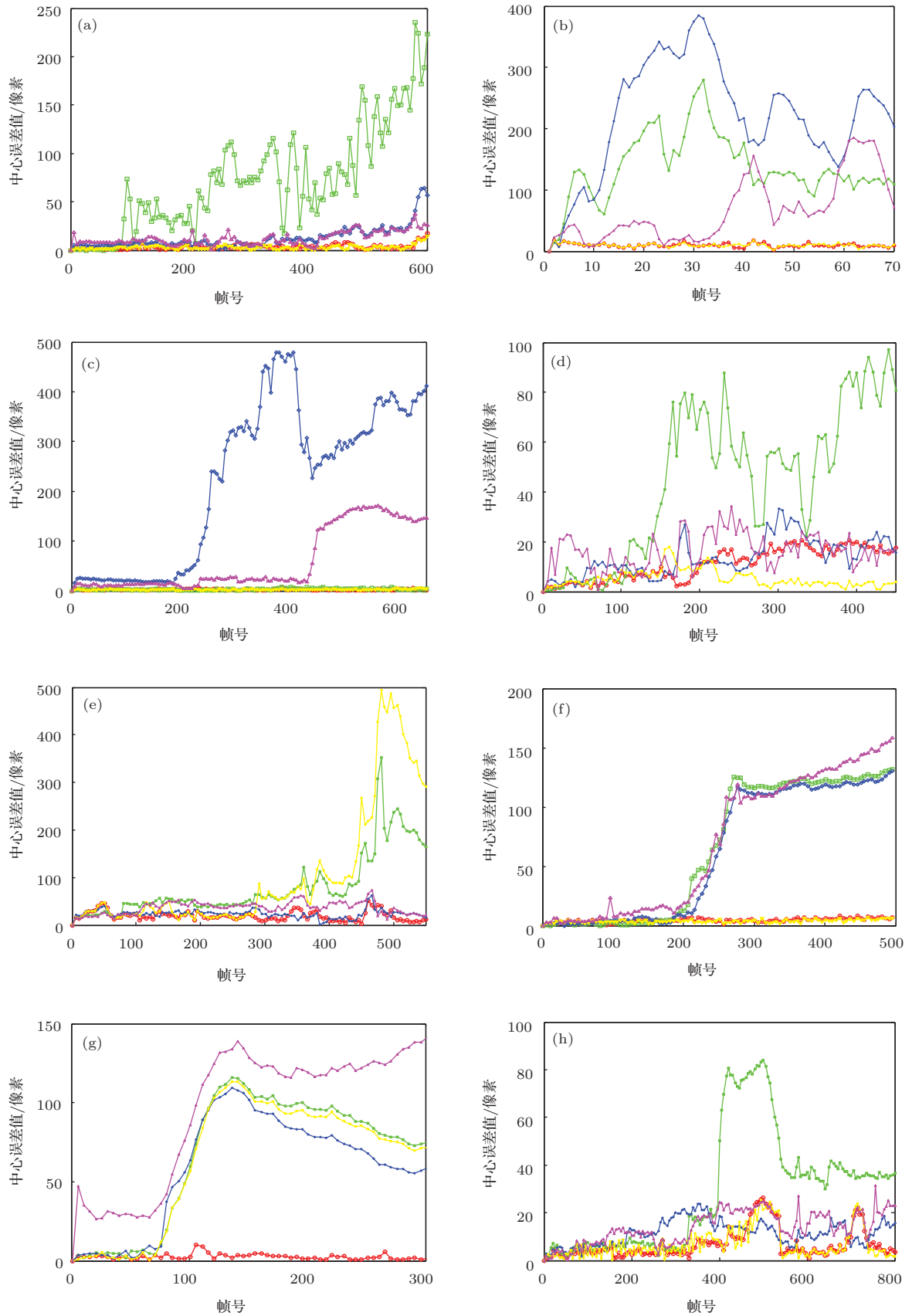


图3 (网刊彩色) 5种跟踪算法的中心定位误差曲线 (a) 对应 boy 视频; (b) 对应 Deer 视频; (c) 对应 Car4 视频; (d) 对应 DavidIndoor 视频; (e) 对应 Dudek 视频; (f) 对应 Caviar2 视频; (g) 对应 Caviar3 视频; (h) 对应 Occlusion2 视频

5 结 论

本文通过对传统的正向稀疏表示和反向稀疏表示目标跟踪框架进行分析,提出一种广义的双向稀疏表示目标跟踪模型.该模型通过 L_2 范数损失约束项在正反向稀疏相关系数矩阵之间构建了一个正反向联合稀疏求解框架.与之前的单向稀疏表示模型相比,双向稀疏表示跟踪模型可以更加充分地求解候选样本与模板集之间的稀疏映射关系.同时为了提高计算效率,本文基于APG快速算法以矩阵方式推导了双向稀疏表示模型的求解框架.实验数据表明,在目标发生旋转、遮挡、尺度变化以及快速运动等复杂场景中,相比于其他跟踪算法(包含单向稀疏表示跟踪算法),本文算法具有更好的鲁棒性.

参考文献

- [1] Gao W, Tang Y, Zhu M 2014 *Acta Phys. Sin.* **63** 094204 (in Chinese) [高文, 汤洋, 朱明 2014 物理学报 **63** 094204]
- [2] Donoho D L 2006 *IEEE Trans. Inform. Theory* **52** 1289
- [3] Wright J, Yang A Y, Ganesh A, Sastry S S, Ma Y 2009 *IEEE Trans. Pattern Anal. Mach. Intell.* **31** 210
- [4] Deng C Z, Tian W, Chen P, Wang S Q, Zhu H S, Hu S F 2014 *Acta Phys. Sin.* **63** 044202 (in Chinese) [邓承志, 田伟, 陈盼, 汪胜前, 朱华生, 胡赛凤 2014 物理学报 **63** 044202]
- [5] Song C X, Ma K, Qin C, Xiao P 2013 *Acta Phys. Sin.* **62** 040702 (in Chinese) [宋长新, 马克, 秦川, 肖鹏 2013 物理学报 **62** 040702]
- [6] Mei X, Ling H B 2011 *IEEE Trans. Pattern Anal. Mach. Intell.* **33** 2259
- [7] Bao C L, Wu Y, Ling H B, Ji H 2012 *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition* Providence, USA, June 16–21, 2012 p1830
- [8] Zhong W, Lu H C, Yang M H 2012 *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition* Providence, USA, June 16–21, 2012 p1838
- [9] Liu H P, Sun F C 2010 *Proceedings of IEEE 20th Conference on Pattern Recognition* Istanbul, Turkey, August 23–26, 2010 p1702
- [10] Zhuang B H, Lu H C, Xiao Z Y, Wang D 2014 *IEEE Trans. Image Proc.* **23** 1872
- [11] Li X, Hu W M, Shen C H, Zhang Z F, Dick A 2013 *ACM Trans. Intell. Syst. Technol.* **4** 58
- [12] Toh K C, Yun S 2010 *Pac. J. Optim.* **6** 615
- [13] Zhang K H, Zhang L, Yang M H 2012 *Proceedings of European Conference on Computer Vision* Berlin, Germany, October 7–13, 2012 p864
- [14] Babenko B, Yang M H, Belongie S 2011 *IEEE Trans. Pattern Anal. Mach. Intell.* **33** 1619
- [15] Wu Y, Lim J, Yang M H 2013 *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition* Portland, USA, June 23–28, 2013 p2411

Robust visual tracking algorithm based on bidirectional sparse representation*

Wang Bao-Xian Zhao Bao-Jun[†] Tang Lin-Bo Wang Shui-Gen Wu Jing-Hui

(School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China)

(Received 17 June 2014; revised manuscript received 8 July 2014)

Abstract

At present the visual tracking model based on sparse representation is mainly divided into two types: one is to use the template set to reconstruct candidate samples, which is called forward model; the other is to project the template set into a candidate space, which is called reverse model. What the two models have in common is to compute the sparse correlation coefficient matrix of candidate sample and template set. Based on this, the paper establishes a bidirectional cooperative sparse representation tracking model. Using L_2 -norm constraint item, the forward and reverse sparse correlation matrix coefficients could be uniformly convergent. In comparison to conventional unidirectional sparse tracking model, bidirectional sparse tracking model could fully excavate the sparse mapping relation of the whole candidate sample and template set. And the candidate that scores highest in the sparse mapping table for the positive and negative templates is the tracking result. Based on the accelerated proximal gradient fast method, the paper derives the optimum solution (in matrix form) of bidirectional sparse tracking model. As a result, it allows the candidates and templates to be calculated in parallel, which can improve the calculation efficiency to some extent. Numerical examples show that the proposed tracking algorithm has certain priority over against the conventional unidirectional sparse tracking methods.

Keywords: sparse representation, visual tracking, Bayesian framework

PACS: 42.30.Va, 07.05.Pj, 42.30.Tz, 02.50.-r

DOI: 10.7498/aps.63.234201

* Project supported by the National High Technology Research and Development Program of China (Grant No. 2012AA8012011C).

[†] Corresponding author. E-mail: zbj@bit.edu.cn