

基于稀疏度约束与动态组结构 稀疏编码的鲁棒视觉跟踪

袁广林¹, 薛模根²

(1. 解放军陆军军官学院十一系, 安徽合肥 230031; 2. 解放军陆军军官学院科研部, 安徽合肥 230031)

摘 要: 目标编码系数的稀疏性使得 L1 跟踪成为解决遮挡目标跟踪的有效方法之一, 但是现有稀疏编码算法没有利用 L1 跟踪中编码系数的特殊稀疏结构. 本文基于目标模板系数稀疏度约束要求和小模板系数的空间连续性结构, 利用块坐标优化原理提出一种两阶段稀疏编码算法用于视觉跟踪. 在第一阶段, 该算法利用正交匹配追踪求解具有约束稀疏度的目标模板系数, 在第二阶段, 该算法利用动态组稀疏编码求解具有空间连续性的小模板系数. 在粒子滤波框架下, 利用提出的稀疏编码算法实现了鲁棒的视觉跟踪. 实验结果表明本文提出的跟踪方法比现有跟踪方法具有更强的鲁棒性和较高的跟踪精度.

关键词: L1 跟踪; 稀疏编码; 约束稀疏度; 空间连续性结构

中图分类号: TP391.4 **文献标识码:** A **文章编号:** 0372-2112 (2015)08-1499-07

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

DOI: 10.3969/j.issn.0372-2112.2015.08.005

Sparsity-Constrained and Dynamic Group Structured Sparse Coding for Robust Visual Tracking

YUAN Guang-lin¹, XUE Mo-gen²

(1. Eleventh Department, Army Officer Academy of PLA, Hefei, Anhui 230031, China;

2. Department of Scientific Research, Army Officer Academy of PLA, Hefei, Anhui 230031, China)

Abstract: L1 tracker is one of the most effective methods in dealing with the occlusions for sparseness of coding coefficients of objects. However, the existing sparse coding algorithms do not use special sparse structure of coding coefficients in L1 tracker. In this paper, we propose a two-stage sparse coding algorithm for visual tracking based on constrained sparsity of target template coefficients and spatial continuity structure of trivial template coefficients with block coordinate optimization theory. At the first stage, the algorithm solves sparsity-constrained coding coefficients on target template set using orthogonal matching pursuit. At the second stage, the algorithm finds sparse coding coefficients with spatial continuity on trivial template set via dynamic group sparse coding. Robust visual tracking is achieved using the proposed sparse coding algorithm in particle filter framework. The experimental results demonstrate that the proposed tracking method has better robustness and higher precision than the state-of-the-art trackers.

Key words: L1 tracker; sparse coding; sparsity-constrained; spatial continuity structure

1 引言

视觉跟踪是计算机视觉中的基本问题之一, 其目的是利用图像序列估计目标的状态参数. 视觉跟踪在运动分析、视频监控、车辆导航和人机交互等方面具有重要的应用价值. 尽管视觉跟踪领域中的研究者已经提出了一些跟踪方法^[1-3], 但是鲁棒快速的视觉跟踪方

法还没有解决. 视觉跟踪面临着目标遮挡、目标变化、复杂背景和快速性要求等难题, 它仍然是计算机视觉领域研究的热点之一.

近年来, 以 D. Donoho 等人^[4]提出的压缩感知理论为基础, 发展起来的稀疏编码理论在计算机视觉中得到了广泛应用. 受到稀疏编码在人脸识别^[5]中应用的启发, 2009年 Mei 和 Ling^[6]首次将稀疏编码应用到视

觉跟踪中,提出了 L1 跟踪. L1 跟踪较好地解决了目标遮挡问题,因而受到了广泛关注. 为了利用目标模板的时间聚类以及鉴别性特征的空间聚类性, Liu 等人^[7]将动态组稀疏编码引入到视觉跟踪中,提高了 L1 跟踪的速度和鲁棒性. Mei 等人^[8]提出一种最小化误差界采样策略,减少了 L1 最小化的求解次数,从而提高 L1 跟踪的速度. Bao 等人^[9]提出在 L1 最小化模型上增加小模板系数的 L2 范数正则化项,并设计了其快速求解算法,进一步提升了 L1 跟踪的鲁棒性与速度. 根据粒子在同一模板集上稀疏编码系数的相似性, Zhang 等人^[10]提出利用多任务稀疏学习联合求解粒子的稀疏编码系数,提高了 L1 跟踪的速度和鲁棒性. Jia 等人^[11]提出一种结构化的局部稀疏编码模型,该模型的编码系数包含了目标的局部与空间信息,进一步提高了 L1 跟踪的精度. 姜明新等人^[12]提出了一种基于最大似然估计和 L2 范数的视觉跟踪方法,该方法将视觉跟踪等价于 L2 范数正则化加权最小二乘问题. 与 L1 跟踪相比,该方法具有较强的鲁棒性和较高的速度. 原始 L1 跟踪及其改进方法^[6-10]利用灰度模板构建模板字典,并用跟踪结果替换模板实现字典更新. 与这些方法不同,文献^[13]用正交完备基设计模板字典并利用增量主分量分析更新模板字典,提升了目标跟踪的鲁棒性和速度.

综上所述,稀疏编码算法对视觉跟踪的性能至关重要. 尽管一些学者已经提出一些稀疏编码算法用于视觉跟踪,但是 L1 跟踪对离群模板(目标模板中的离群数据)较敏感,一旦目标模板中引入离群数据会导致跟踪失败,这一问题并没有解决. 另一方面, L1 跟踪中的小模板系数具有空间连续性稀疏结构,但是现有 L1 跟踪方法并未利用这一特殊的稀疏结构. 基于这两个方面,本文利用块坐标优化原理提出一种称为“稀疏度约束与动态组结构稀疏编码”的两阶段稀疏编码算法用于视觉跟踪. 在粒子滤波框架下,利用提出的稀疏编码算法实现了目标跟踪方法. 实验结果表明:本文提出的稀疏编码算法对视觉跟踪问题实现了鲁棒性与跟踪精度提高的预期效果.

2 L1 跟踪方法

L1 跟踪^[6]利用稀疏编码求解粒子的观测似然,以粒子滤波为框架跟踪目标.

2.1 粒子滤波

粒子滤波包括预测和更新两个过程. 假设 $\mathbf{y}_{1:t-1} = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_{t-1}\}$ 为 1 到 $t-1$ 时刻的图像观测,则预测过程为:

$$p(\mathbf{x}_t | \mathbf{y}_{1:t-1}) = \int p(\mathbf{x}_t | \mathbf{x}_{t-1}) p(\mathbf{x}_{t-1} | \mathbf{y}_{1:t-1}) d\mathbf{x}_{t-1} \quad (1)$$

其中, \mathbf{x}_t 表示 t 时刻的目标状态, $p(\mathbf{x}_k | \mathbf{x}_{k-1})$ 是运动模型. 在 t 时刻,观测 \mathbf{y}_t 可用,执行如下更新过程:

$$p(\mathbf{x}_t | \mathbf{y}_{1:t}) = \frac{p(\mathbf{y}_t | \mathbf{x}_t) p(\mathbf{x}_t | \mathbf{y}_{1:t-1})}{p(\mathbf{y}_t | \mathbf{y}_{1:t-1})} \quad (2)$$

其中, $p(\mathbf{y}_t | \mathbf{x}_t)$ 为观测似然模型. 式(2)中后验概率 $p(\mathbf{x}_t | \mathbf{y}_{1:t})$ 用粒子集 $\{\mathbf{x}_t^i, w_t^i\}_{i=1}^N$ 近似,其中, N 是粒子数, \mathbf{x}_t^i 和 w_t^i 分别是 t 时刻第 i 个粒子的状态和权值. 通过重要性分布 $q(\mathbf{x}_t | \mathbf{x}_{t-1}, \mathbf{y}_{1:t})$ 采样粒子,利用式(3)更新粒子权值:

$$w_t^i = w_{t-1}^i \frac{p(\mathbf{y}_t | \mathbf{x}_t^i) p(\mathbf{x}_t^i | \mathbf{x}_{t-1}^i)}{q(\mathbf{x}_t | \mathbf{x}_{t-1}^i, \mathbf{y}_{1:t})} \quad (3)$$

一般取 $q(\mathbf{x}_t | \mathbf{x}_{t-1}, \mathbf{y}_{1:t}) = p(\mathbf{x}_t | \mathbf{x}_{t-1})$, 这样权值更新可表示为: $w_t^i = w_{t-1}^i p(\mathbf{y}_t | \mathbf{x}_t^i)$.

2.2 稀疏编码

稀疏编码在 L1 跟踪中的作用是求解粒子 \mathbf{x}_t 的观测似然 $p(\mathbf{y}_t | \mathbf{x}_t)$. 假设 t 时刻的目标模板为 $\mathbf{T}_t = [\mathbf{t}_t^1, \mathbf{t}_t^2, \dots, \mathbf{t}_t^n] \in \mathbb{R}^{d \times n}$, 粒子状态集为 $\mathbf{X}_t = \{\mathbf{x}_t^1, \mathbf{x}_t^2, \dots, \mathbf{x}_t^N\}$, 与 \mathbf{X}_t 对应的粒子的图像观测为 $\mathbf{Y}_t = \{\mathbf{y}_t^1, \mathbf{y}_t^2, \dots, \mathbf{y}_t^N\}$, 则任一图像观测 \mathbf{y}_t^i 表示如下:

$$\mathbf{y}_t^i = \mathbf{T}_t \mathbf{a}_t^i + \mathbf{I} \mathbf{e}_t^i = [\mathbf{T}_t, \mathbf{I}] \begin{bmatrix} \mathbf{a}_t^i \\ \mathbf{e}_t^i \end{bmatrix} = \mathbf{D}_t \mathbf{c}_t^i \quad (4)$$

其中 $\mathbf{a}_t^i = [a_t^{i,1}, a_t^{i,2}, \dots, a_t^{i,n}]^T \in \mathbb{R}^n$ 是图像观测 \mathbf{y}_t^i 在目标模板 \mathbf{T}_t 上的编码系数; $\mathbf{I} \in \mathbb{R}^{d \times d}$ 是一个单位阵,称为小模板; $\mathbf{e}_t^i = [e_t^{i,1}, e_t^{i,2}, \dots, e_t^{i,d}]^T \in \mathbb{R}^d$ 是图像观测 \mathbf{y}_t^i 在单位阵 \mathbf{I} 上的编码系数,称为小模板系数; $\mathbf{c}_t^i = [\mathbf{a}_t^i, \mathbf{e}_t^i]^T$ 是图像观测 \mathbf{y}_t^i 在模板字典 $\mathbf{D}_t = [\mathbf{T}_t, \mathbf{I}] \in \mathbb{R}^{d \times (n+d)}$ 上的编码系数. 图像观测 \mathbf{y}_t^i 的编码系数 \mathbf{c}_t^i 用式(5)求解.

$$\min_{\mathbf{c}_t^i} \frac{1}{2} \|\mathbf{y}_t^i - \mathbf{D}_t \mathbf{c}_t^i\|_2^2 + \lambda \|\mathbf{c}_t^i\|_1 \quad (5)$$

其中 $\|\cdot\|_1$ 和 $\|\cdot\|_2$ 分别表示 L1 范数和 L2 范数, λ 是正则化参数. L1 跟踪用式(5)求解候选目标 \mathbf{x}_t^i 的编码系数 \mathbf{c}_t^i , 用式(6)计算 \mathbf{x}_t^i 的观测似然.

$$p(\mathbf{y}_t | \mathbf{x}_t^i) = \frac{1}{\Gamma} \exp\{-\alpha \|\mathbf{y}_t^i - \mathbf{T}_t \mathbf{a}_t^i\|_2^2\} \quad (6)$$

其中 α 是高斯核尺度参数, Γ 是归一化常量. 在 t 时刻, L1 跟踪用最大后验概率准则估计目标的状态 $\hat{\mathbf{x}}_t$, 如式(7)所示.

$$\hat{\mathbf{x}}_t = \arg \max_{\mathbf{x}_t \in \mathbf{X}_t} p(\mathbf{y}_t | \mathbf{x}_t^i), i = 1, 2, \dots, N \quad (7)$$

2.3 L1 跟踪的鲁棒性分析

L1 跟踪的本质是子空间跟踪,子空间跟踪有传统子空间跟踪^[1]和 L1 跟踪^[6]两类. 传统子空间跟踪以 $\mathbf{y} = \mathbf{T}\mathbf{a} + \mathbf{n}$ 为产生式模型,其中 \mathbf{n} 是高斯噪声,这与视觉跟踪的情况不符,导致它对目标遮挡的鲁棒性较差. 在

视觉跟踪中,由于目标遮挡使得观测噪声符合拉普拉斯分布.为此,文献[6]以 $\mathbf{y}=\mathbf{Ta}+\mathbf{e}$ 为产生式模型提出了 L1 跟踪,其中 \mathbf{e} 是拉普拉斯噪声. $\mathbf{y}=\mathbf{Ta}+\mathbf{e}$ 可以写成 $\mathbf{y}=\mathbf{Dc}$,其中 $\mathbf{D}=[\mathbf{T},\mathbf{I}],\mathbf{c}=[\mathbf{a};\mathbf{e}]$. L1 跟踪假设遮挡仅占目标的小部分区域,所以好的候选目标其噪声向量 \mathbf{e} 仅有少部分非零项,即 \mathbf{e} 具有稀疏性,因此 $\mathbf{y}=\mathbf{Dc}$ 具有稀疏解.据此,L1 跟踪用式(5)求解候选目标的稀疏编码系数 \mathbf{c} ,以后选目标在目标模板 \mathbf{T} 上的重建误差 $\mathbf{y}-\mathbf{Ta}$ 计算观测似然跟踪目标.正是上述原因,使得 L1 跟踪对目标遮挡具有鲁棒性.

3 稀疏度约束与动态组稀疏编码

由 2.3 节的分析可知,L1 跟踪对目标遮挡具有鲁棒性的原因是式(5)中对编码系数 \mathbf{c} 的稀疏性约束.但是编码系数的稀疏性约束使得 L1 跟踪对目标模板 \mathbf{T} 中的离群数据较敏感,当目标模板中引入离群数据时,易于引起跟踪失败^[7].我们研究发现,L1 跟踪对模板中离群数据敏感是候选目标在 \mathbf{T} 上的编码系数的稀疏度过小造成的.编码系数的稀疏度过小使得候选目标的重建误差仅取决于 \mathbf{T} 中少数几个模板,甚至是某一个模板. L1 跟踪用候选目标在 \mathbf{T} 上的重建误差计算观测似然跟踪目标,所以当目标模板中引入背景样本时易产生跟踪失败.另一方面,由 2.3 节的分析可知,为了保证对遮挡目标跟踪的鲁棒性,对小模板系数 \mathbf{e} 进行 L1 范数正则化是必须的.由于目标遮挡使得小模板系数具有不可预知的空间连续性结构^[14],即小模板系数具有动态组稀疏结构,然而现有 L1 跟踪及其改进方法并没有利用小模板系数这一性质.基于这两个方面,本文提出稀疏度约束与动态组稀疏编码算法求解目标模板系数和小模板系数.为了方便描述,略去式(5)中的上下标符号并将其重写为式(8).

$$\min_{\mathbf{a},\mathbf{e}} \frac{1}{2} \|\mathbf{y}-\mathbf{Ta}-\mathbf{e}\|_2^2 + \lambda_1 \|\mathbf{a}\|_1 + \lambda_2 \|\mathbf{e}\|_1 \quad (8)$$

其中 λ_1 和 λ_2 是正则化参数,其它符号与式(5)相同.

引理 1 给定最优解 \mathbf{e}_{opt} ,则可用正交匹配追踪算法求解最优解 \mathbf{a}_{opt} ,且 \mathbf{a}_{opt} 的稀疏度可控.

证明 假如 \mathbf{e}_{opt} 已知,则优化问题式(8)可表示为 $\min_{\mathbf{a}} \frac{1}{2} \|\tilde{\mathbf{y}}-\mathbf{Ta}\|_2^2 + \lambda_1 \|\mathbf{a}\|_1$,其中 $\tilde{\mathbf{y}}=\mathbf{y}-\mathbf{e}$.此问题等价于 L0 最小化问题: $\min_{\mathbf{a}} \|\mathbf{a}\|_0$, s. t. $\|\tilde{\mathbf{y}}-\mathbf{Ta}\|_2 < \varepsilon$.由文献[15]知,其最优解 \mathbf{a}_{opt} 可用算法 1 正交匹配追踪算法求解.由于正交匹配追踪算法可以指定编码系数的稀疏度,所以 \mathbf{a}_{opt} 的稀疏度可控.由算法 1 可知,在目标模板确定时,可以通过设置算法 1 中参数 k 控制候选目标在目标模板上的编码稀疏度.

算法 1 中 \mathbf{M}_{Ω} 表示由支集索引 Ω 中元素指示的编码矩阵 \mathbf{M} 中的列向量形成的矩阵,即假设 $\Omega = \{1, 2, \dots, l\}$,则 $\mathbf{M}_{\Omega} = [\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_l]$.算法 2 中 \mathbf{M}_{Ω} 具有相同的含义.

引理 2 给定 \mathbf{a}_{opt} ,则可用动态组稀疏编码算法求解 \mathbf{e}_{opt} ,且 \mathbf{e}_{opt} 具有动态组稀疏结构.

证明 假如 \mathbf{a}_{opt} 已知,则优化问题式(8)可表示为 $\min_{\mathbf{e}} \frac{1}{2} \|\hat{\mathbf{y}}-\mathbf{Ie}\|_2^2 + \lambda_2 \|\mathbf{e}\|_1$,其中 $\hat{\mathbf{y}}=\mathbf{y}-\mathbf{Ta}_{\text{opt}}$.此问题等价于 L0 最小化问题: $\min_{\mathbf{e}} \|\mathbf{e}\|_0$, s. t. $\|\hat{\mathbf{y}}-\mathbf{Ie}\|_2 < \varepsilon$.考虑到 \mathbf{e} 的动态组稀疏结构,由文献[16]知,其最优解 \mathbf{e}_{opt} 可用算法 2 动态组稀疏编码算法求解.由算法 2 可知,动态组稀疏编码算法用外循环动态设置稀疏度 k ,用内循环求解稀疏度为 k 的组稀疏编码,即用算法 2 求解得到的 \mathbf{e}_{opt} 具有动态组稀疏结构.

算法 1 正交匹配追踪算法

输入:观测向量 $\mathbf{u} \in \mathbb{R}^m$,编码矩阵 $\mathbf{M} = [\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_n] \in \mathbb{R}^{m \times n}$,稀疏度 k

输出:编码系数 $\mathbf{c} \in \mathbb{R}^n$

步骤 1. 初始化余量 $\mathbf{r}_0 = \mathbf{u}$,支集索引 $\Omega = \emptyset$

for $i = 1$ to k do

步骤 2. $\kappa = \arg \max_j \langle \mathbf{r}_{i-1}, \mathbf{m}_j \rangle, j = 1, \dots, n$;

步骤 3. $\Omega = \Omega \cup \{\kappa\}$;

步骤 5. $\mathbf{c}_i = \arg \min_{\mathbf{c}} \|\mathbf{M}_{\Omega} \mathbf{c} - \mathbf{u}\|_2$;

步骤 6. $\mathbf{r}_i = \mathbf{u} - \mathbf{M}_{\Omega} \mathbf{c}_i$;

end for

步骤 7. $\mathbf{c} = \mathbf{c}_k$.

算法 2 动态组稀疏编码算法

输入:观测向量 $\mathbf{u} \in \mathbb{R}^m$,编码矩阵 $\mathbf{M} = [\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_n] \in \mathbb{R}^{m \times n}$,稀疏度范围 $[k_1, k_2]$

输出:编码系数 $\mathbf{c} \in \mathbb{R}^n$

步骤 1. 初始化余量 $\mathbf{r}_0 = \mathbf{u}$,支集索引 $\Omega = \emptyset$,稀疏度 $k = k_1, i = 0, f_0 = 0$;

repeat

步骤 2. $i = i + 1$;

步骤 3. $j = 0$;

repeat

步骤 4. $j = j + 1$;

步骤 5. $\mathbf{v} = \mathbf{M}^T \mathbf{r}_j$;

步骤 6. $\Omega_{\mathbf{v}} = \text{DGS}(\mathbf{v}, k)$;

步骤 7. $\Omega = \Omega \cup \Omega_{\mathbf{v}}$;

步骤 8. $\mathbf{b} = \arg \min_{\mathbf{x}} \|\mathbf{M}_{\Omega} \mathbf{x} - \mathbf{u}\|_2$;

步骤 9. $\Omega_{\mathbf{b}} = \text{DGS}(\mathbf{b}, k)$;

步骤 10. $\Omega = \Omega_{\mathbf{b}}$;

步骤 11. $\mathbf{c}_j = \arg \min_{\mathbf{c}} \|\mathbf{M}_{\Omega} \mathbf{c} - \mathbf{u}\|_2$;

步骤 12. $\mathbf{r}_j = \mathbf{u} - \mathbf{M}_{\Omega} \mathbf{c}_j$;

until $\|\mathbf{r}_j\|_2 \geq \|\mathbf{r}_{j-1}\|_2$

步骤 13. $\mathbf{c}_i = \mathbf{c}_j$

步骤 14. $k = k + \Delta k$;

步骤 15. $f_i = 0.5 \|Mc_i - u\|_2 + \|c_i\|_1$;

until $\text{abs}(f_i - f_{i-1}) \leq 0.001$

步骤 16. $x = x_i$

算法 3 DGS 算法

输入: 信号 $x \in \mathbb{R}^n$, 稀疏度 k

输出: x 的支集 $\text{supp}\{x, k\}$

步骤 1. 邻域权重 $w = [0.5, 0.5, 0.5, 0.5]$, 其中 $0.5 = [0.5, 0.5, \dots, 0.5]^T$

步骤 2. 计算 x 中每一元素的 4 邻域索引得到 $N_x \in \mathbb{R}^{n \times 4}$

for $i = 1$ to n do

步骤 3. 计算 $z(i) = x^2(i) + \sum_{t=1}^4 w^2(i, t) N_x^2(i, t)$;

end for

步骤 4. 令 $\text{supp}\{x, k\}$ 为 $z(i)$ 中前 k 个最大值对应的索引值.

由引理 1 和引理 2, 并结合块坐标优化原理^[17]可知, 式(8)可分两个步骤求解: 首先固定小模板系数 e , 求解目标模板系数 a ; 然后固定小模板系数 a , 求解目标模板系数 e . 这样循环迭代这两个步骤直至算法收敛. 假设 $OMP(\cdot, \cdot)$ 表示算法 1 正交匹配追踪算法, $AdaDGS(\cdot, \cdot)$ 表示算法 2 动态组稀疏编码算法, 则本文提出的稀疏度约束与动态组稀疏编码算法如算法 4 所示.

算法 4 稀疏度约束与动态组稀疏编码算法

输入: 候选目标 y , 目标模板 T , 稀疏度 k , 稀疏度范围 $[k_1, k_2]$

输出: a_{opt} 和 e_{opt}

步骤 1. 初始化 $e_0 = 0, i = 0, f_0 = 0, \text{maxLN} = 20$;

repeat:

步骤 2. $i = i + 1$;

步骤 3. $\hat{y} = y - e_{i-1}$;

步骤 4. $a_i = OMP(T, \hat{y}, k)$;

步骤 5. $\hat{y} = y - Ta_i$;

步骤 6. $e_i = AdaDGS(T, \hat{y}, k_1, k_2)$;

步骤 7. $f_i = 0.5 \|y - Ta_i - e_i\|_2 + \|a_i\|_1 + \|e_i\|_1$

until $\text{abs}(f_i - f_{i-1}) \leq 0.001$ 或者 $i \geq \text{maxLN}$

步骤 7. 令 $a_{\text{opt}} = a_i$ 和 $e_{\text{opt}} = e_i$.

由于一次跟踪要用算法 4 求解多个粒子的编码系数, 所以算法 4 的收敛条件直接影响到目标跟踪的精度和实时性. 本文用连续两次迭代目标函数的差值和最大循环次数作为其收敛条件. 实验发现循环迭代 5 次左右算法 4 即可收敛. 第 5 节实验结果表明, 算法 4 的收敛条件能够保证较高的跟踪精度和实时性.

4 目标跟踪方法

本文利用粒子滤波跟踪目标, 主要包括运动模型和观测似然模型两个组件, 其中运动模型如下:

$$p(x_i | x_{i-1}) = N(x_i; x_{i-1}, \Psi) \quad (9)$$

其中 $x_i = (x_i, y_i, w_i, h_i, \theta_i)$ 是目标状态, x_i, y_i, w_i, h_i 和 θ_i 分别是 t 时刻目标区域中心的 x 坐标、 y 坐标、目标的宽度、高度和倾斜角, Ψ 是对角矩阵, 其对角上的元素是状态分量的方差. 对于任一粒子 x_i^j , 本文首先用算法 4 求解式(10)得到其图像观测 y_i^j 的编码系数 e_i^j 和 a_i^j ,

$$\min_{a_i^j, e_i^j} \frac{1}{2} \|y_i^j - T a_i^j - e_i^j\|_2^2 + \lambda_1 \|a_i^j\|_1 + \lambda_2 \|e_i^j\|_1 \quad (10)$$

然后用式(11)计算其观测似然.

$$p(y_i^j | x_i^j) = \frac{1}{\Gamma} \exp\left(-\alpha \left(\frac{1}{2} \|y_i^j - T a_i^j - e_i^j\|_2^2 + \lambda_2 \|e_i^j\|_1\right)\right) \quad (11)$$

其中 α 是高斯核尺度参数, Γ 是归一化常量.

利用粒子滤波将运动模型和观测似然相结合, 建立基于稀疏度约束与动态组结构稀疏编码的跟踪方法, 如方法 1 所示.

方法 1 基于稀疏度约束与动态组结构稀疏编码的跟踪方法

输入: 目标初始状态 x_0 和目标序列图像

输出: t 时刻目标状态 \hat{x}_t

步骤 1. 初始化: 初始化目标模板 T_0 ;

步骤 2. 采样粒子: 用运动模型式(9)采样粒子 x_i^j , 并得到其图像观测 $y_i^j, i = 1, 2, \dots, N$;

步骤 3. 计算观测似然: 用算法 4 求解粒子 y_i^j 的编码系数 a_i^j 和 e_i^j , 并用式(11)计算粒子 x_i^j 观测似然, $i = 1, 2, \dots, N$;

步骤 4. 估计目标状态: 用最大后验概率准则估计目标状态 \hat{x}_t , 即: $\hat{x}_t = \arg \max_{x_t^j} p(x_t^j | y_{1:t}) = \arg \max_{x_t^j} p(y_t^j | x_t^j) p(x_t^j | x_{t-1}^j), i = 1, 2, \dots, N$;

步骤 5. 更新模板: 用文献[6]中的算法 1 更新目标模板 T_i ;

步骤 6. 如果已到最后一帧, 则结束跟踪, 否则转至步骤 2.

方法 1 采用与 IVT 跟踪^[1]相同的方法初始化目标模板 T_0 , 即前 n 帧利用候选目标与当前模板均值向量的差计算观测似然跟踪目标, 利用前 n 帧的跟踪结果初始化目标模板 T_0 . 在目标跟踪过程中, 用文献[6]中的算法 1 更新目标模板 T_i .

5 实验结果与分析

5.1 实验说明

以 Matlab R2010a 为开发工具, 在 Inter(R) Forth-Core 2. 50GHz CPU, 4G 内存的台式机上, 实现了本文提出的跟踪方法. 用 4 个具有挑战性的图像序列^[18]对提出的跟踪方法进行了实验验证, 并与 IVT 跟踪^[1]、LIAPG 跟踪^[9]、MTT 跟踪^[10]和 LSS 跟踪^[13]进行了比较. 实验中, 目标的真实状态和其它 4 个跟踪方法在这 4 个测试序列上的跟踪结果由标准测试数据库得到^[18]. 本文跟踪方法的粒子数为 600, 正则化参数 λ_1 和

λ_2 的值均取 1, 稀疏度 k 为 8, 模板数是 16, faceocc2、david、football 和 jumping 序列的模板大小分别是 $22 \times$

$26, 16 \times 22, 20 \times 20$ 和 16×22 , 每隔 5 帧进行一次模板更新.

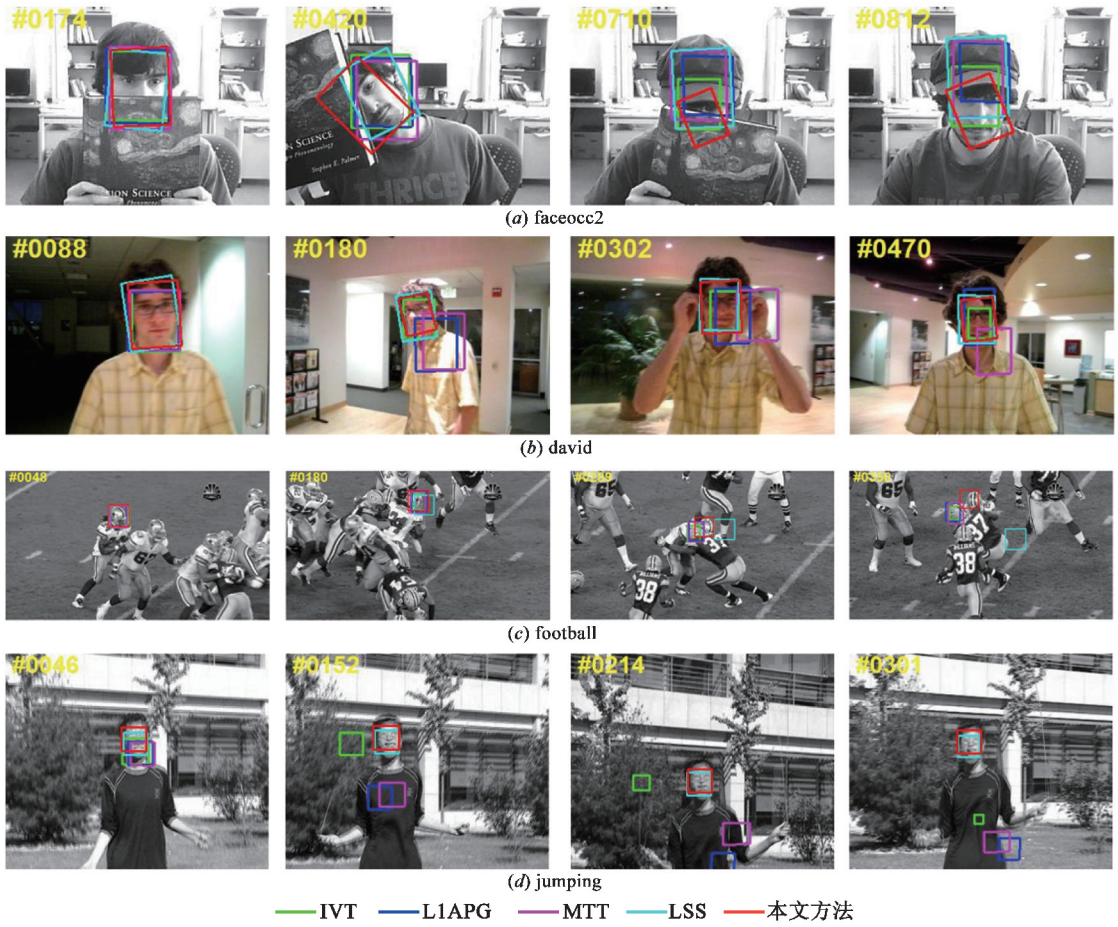


图1 目标跟踪结果

5.2 实验结果

图 1(a) 给出了 5 个跟踪方法对遮挡目标的跟踪结果. 由图 1(a) 可以看出: 对于 faceocc2 序列, 本文跟踪方法和 LSS 跟踪的跟踪结果较好, 尤其是目标同时存在遮挡和 in-plane 旋转时, 本文跟踪方法的结果最优. 图 1(b) 给出了 5 个跟踪方法对光照变化目标的跟踪结果. 由图 1(b) 可以看出: 对于存在光照变化、out-plane 旋转和尺度变化等诸多挑战的 david 序列, 本文跟踪方法和 IVT 跟踪的结果优于其它 3 种跟踪方法. 图 2(c) 给出了 5 个跟踪方法对复杂背景下目标的跟踪结果. football 序列的背景复杂, 同时也存在 out-plane 旋转、遮挡和 in-plane 旋转等挑战, 由图 1(c) 可以看出: 对于 football 序列, 本文跟踪方法的结果最优. 图 1(d) 给出了 5 个跟踪方法对快速运动目标的跟踪结果. 由图 1(d) 可以看出: 对于目标的快速运动而导致目标模糊的 jumping 序列, 本文跟踪方法和 LSS 跟踪的结果最优.

假设目标的真实矩形区域和跟踪结果矩形区域分别是 r_a 和 r_t , 则中心位置误差定义为: r_a 和 r_t 中心位置之间的欧式距离 (单位为像素). 跟踪精度定义为: 中心位置

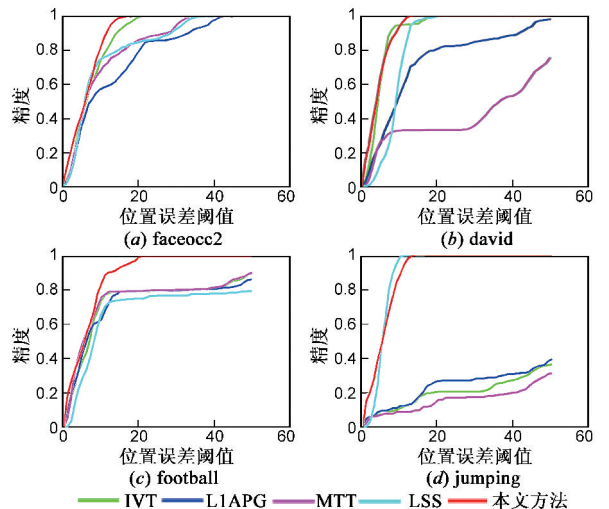


图2 跟踪精度曲线

本文利用跟踪精度定量评价跟踪方法的性能^[18].

误差小于给定位置误差阈值的图像帧数与序列图像总帧数的比值. 给定一系列位置误差阈值即可得到跟踪精度曲线. 图 2 给出了 5 种跟踪方法对 4 组图像序列的跟踪精度曲线, 其中, 中心位置误差阈值为 0~50 像素. 由图 2 可以看出: 本文提出的跟踪方法优于其它 4 个跟踪方法.

5.3 鲁棒性分析

由上述实验可知: 与 L1 跟踪及其改进方法相比, 本文方法具有较强的鲁棒性与较高的跟踪精度. 其原因在于本文跟踪方法与 L1 跟踪及其改进方法使用的稀疏编码算法不同. 一方面, 本文提出的稀疏编码算法求得的模板系数的稀疏度可控, 这样能够保证模板编码系数的稀疏度不过于稀疏, 这样目标跟踪不依赖于少数几个模板, 而是较多模板共同作用的结果, 即使模板中存在离群数据仍能稳定跟踪目标, 这提高了跟踪方法对模板中离群数据的鲁棒性. 另一方面, 本文提出的稀疏编码算法不但使用了小模板系数的稀疏性, 而且也使用了小模板系数的空间连续性结构, 这样不但保证了跟踪方法对遮挡的鲁棒性, 又进一步提高了跟踪精度.

5.4 复杂性分析

假设 $U \in \mathbf{R}^{d \times n}$ 是 IVT 跟踪和 LSS 跟踪使用的特征基, $T \in \mathbf{R}^{d \times n}$ 是 LIAPG 跟踪、MTT 跟踪和本文方法使用的目标模板, k 是 LSS 回归算法^[13]、APG 算法^[9]和本文算法 1 的迭代次数, N 是这 5 个跟踪方法使用的粒子数. 由文献[1, 9, 10, 13]可知: IVT 跟踪、LIAPG 跟踪、MTT 跟踪和 LSS 跟踪的时间复杂度如表 1 所示. 与其它 4 个跟踪方法一样, 本文跟踪方法的时间复杂度取决于粒子数与计算单个粒子观测似然的时间复杂度. 单个粒子观测似然的时间复杂度由本文算法 4 的时间复杂度确定. 本文算法 4 的最耗时部分是其第 3 步, 其时间复杂度是 $O(kdn)$, 所以本文跟踪方法的时间复杂度是 $O(Nkdn)$. 由上述分析可知: 这 5 个跟踪方法的时间复杂度级别相同, 都是变量多项式的复杂度. 表 1 中的第 3 列给出了在相同软硬件环境下, $d=32 \times 32$, $n=16$ 和 $N=600$ 时, 进行一次跟踪的计算时间. 由表 1 可以看出: 本文跟踪方法的速度比 IVT 跟踪和 LSS 跟踪慢, 但是比 LIAPG 跟踪和 MTT 跟踪的速度快.

表 1 计算复杂度与计算时间

跟踪方法	计算复杂度	计算时间(s)
IVT 跟踪 ^[1]	$O(Ndn)$	0.128
LIAPG 跟踪 ^[9]	$O(Nkdn)$	2.274
MTT 跟踪 ^[10]	$O(Nkdn)$	1.304
LSS 跟踪 ^[13]	$O(Nkdn)$	0.268
本文跟踪	$O(Nkdn)$	0.547

6 结论

本文利用块坐标优化原理提出一种稀疏度约束与动态组稀疏编码算法. 提出的算法包括两个阶段, 第一阶段利用正交匹配追踪求解候选目标的稀疏度约束模板编码系数, 这保证了 L1 跟踪对模板中离群数据的鲁棒性; 第二阶段利用动态组稀疏编码求解候选目标的小模板系数, 这利用了小模板系数的动态组稀疏结构, 进一步提升了 L1 跟踪的精度. 以粒子滤波为框架, 利用提出的稀疏编码算法实现了鲁棒的跟踪方法. 在多个图像序列上的实验结果表明: 提出的跟踪方法在鲁棒性和跟踪精度方面优于现有跟踪方法.

参考文献

- [1] 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.
- [2] Kwon J, Lee K M. Visual tracking decomposition [A]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition [C]. San Francisco: IEEE Computer Society Press, 2010. 1269-1276.
- [3] 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.
- [4] Donoho D. Compressed sensing [J]. IEEE Transactions on Information Theory, 2006, 52(4): 1289-1306.
- [5] 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.
- [6] Xue Mei, Haibin Ling. Robust visual tracking using L1 minimization [A]. Proceedings of IEEE Conference on Computer Vision [C]. Kyoto: IEEE Computer Society Press, 2009. 1436-1443.
- [7] Baiyang Liu, Lin Yang, Junzhou Huang, et al. Robust and fast collaborative tracking with two stage sparse optimization [A]. Proceedings of Conference on European Conference on Computer Vision [C]. Crete: Springer, 2010. Part IV: 624-637.
- [8] Mei X, Ling H B, Wu Y, et al. Minimum error bounded efficient L1 tracker with occlusion detection [A]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition [C]. Colorado: IEEE Computer Society Press, 2011. 1257-1264.
- [9] Chenglong Bao, Yi Wu, Haibin Ling, et al. Real time robust L1 tracker using accelerated proximal gradient approach

- [A]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition[C]. Rhode Island:IEEE Computer Society Press,2012. 1830–1837.
- [10] Tianzhu Zhang, Bernard Ghanem, Si Liu, et al. Robust visual tracking via multi-task sparse learning [J]. International Journal of Computer Vision,2013, 101(2): 367–383.
- [11] Xu Jia, Huchuan Lu, Ming-Hsuan Yang. Visual tracking via adaptive structural local sparse appearance model [A]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition[C]. Rhode Island:IEEE Computer Society Press,2012. 1822–1929.
- [12] 姜明新,王洪玉,王洁,等. 基于 ML 和 L2 范数的视频目标跟踪算法 [J]. 电子学报, 2013, 41(11): 2307–2313.
Jiang Ming-xin, Wang Hong-yu, Wang Jie, et al. Visual object tracking algorithm based on ML and L2 norm[J]. Acta Electronica Sinica, 2013, 41(11): 2307–2313. (in Chinese)
- [13] Dong Wang, Huchuan Lu, Ming-Hsuan Yang. Least soft-threshold squares tracking [A]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition [C]. Portland: IEEE Computer Society Press, 2013. 2371–2378.
- [14] Zihan Zhou, Andrew Wagner, Hossein Mobahi, et al. Face recognition with contiguous occlusion using Markov random fields [A]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition[C]. Kyoto: IEEE Computer Society Press, 2009. 1050–1057.
- [15] Joel A Tropp, Anna C Gilbert. Signal recovery from random measurements via orthogonal matching pursuit [J]. IEEE Transactions on information theory, 2007, 53(12): 4655–4666.
- [16] Junzhou Huang, Xiaolei Huang, Dimitris Metaxas. Learning with dynamic group sparsity [A]. Proceedings of IEEE Conference on Computer Vision [C]. Kyoto: IEEE Computer Society Press, 2009. 64–71.
- [17] Paul Tseng, Sangwoon Yun. A block-coordinate gradient descent method for linearly constrained nonsmooth separable optimization [J]. Optimization Theory Application, 2009, 140: 513–535.
- [18] Yi Wu, Jongwoo Lim, Ming-Hsuan Yang. Online object tracking: a benchmark [A]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition [C]. Portland: IEEE Computer Society Press, 2013. 2411–2418.

作者简介



袁广林 男, 1973 年生于河南周口. 博士, 讲师, 研究方向为图像处理、计算机视觉、机器学习及其应用等.

E-mail: yuanguanglin1008@sina.com



薛模根 男, 1964 年生于安徽合肥. 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.