

## 基于联合模板稀疏表示的目标跟踪方法

张旭东, 陈仲海, 胡良梅, 杨 慧, 董文菁

(合肥工业大学 计算机与信息学院, 合肥 230009)

**摘要:** 针对传统基于稀疏表示的目标跟踪方法中, 当场景中含有与目标相似的背景时容易出现跟踪漂移的问题, 提出一种新的目标跟踪方法. 该方法基于目标的局部二元模式特征, 将目标外观模型同时用原始目标模板与当前帧部分粒子构成的联合模板稀疏表示, 构建一个联合目标函数, 将跟踪问题通过迭代转化为求解最优化问题. 实验结果表明, 所提出跟踪方法在解决遮挡、光照等问题的同时, 对场景中含有与目标相似背景的序列具有较好的跟踪效果.

**关键词:** 目标跟踪; 联合模板; 联合目标函数; 稀疏表示

**中图分类号:** TP391

**文献标志码:** A

## Object tracking method based on sparse representation of joint template

ZHANG Xu-dong, CHEN Zhong-hai, HU Liang-mei, YANG Hui, DONG Wen-jing

(School of Computer and Information, Hefei University of Technology, Hefei 230009, China. Correspondent: ZHANG Xu-dong, E-mail: xudong@hfut.edu.cn)

**Abstract:** When a scene contains a similar background to the object, the traditional tracking method based on sparse representation may produce the drifting problems, a new object tracking method is proposed. Based on the local binary pattern features(LBP) of the object, the object appearance model can sparse representation simultaneously with the original object template and the template built by some particles in the current frame, and build a joint objective function which can solve the tracking issue through an iterative optimization problem. Experimental results show that the track method can well track with occlusion and illumination issues, as well as the scene sequence with a similar background to the object.

**Keywords:** object tracking; joint template; joint object function; sparse representation

### 0 引言

目标跟踪是计算机视觉中的一个重要研究领域, 它不仅可以提供目标的运动状态和轨迹, 也为运动分析、场景理解、行为检测等提供初始数据, 在智能监控、人机交互和机器人视觉导航等领域有着广泛的应用前景<sup>[1]</sup>.

虽然跟踪算法在过去的几十年中有了很大的发展, 但由于受到包括目标自身姿态、形状变化等内部因素, 以及遮挡、光照变化、视角变化、杂波背景等外部因素的影响, 使目标外观模型随时间而变化, 从而使运动目标跟踪的发展仍面临极大的挑战. 目前已有的跟踪算法可以归为产生式方法和判别式方法. 产生式方法将目标跟踪问题看作一个模型匹配的问题, 通过搜索具有最小重建误差的图像区域得到跟踪结果, 如: 基于模板的产生式方法<sup>[2-3]</sup>, 基于子空间模型

产生式方法<sup>[4-6]</sup>, 使用多运动模型、多外观模型对传统粒子滤波进行改进的方法<sup>[7]</sup>和基于稀疏表示的跟踪方法<sup>[8]</sup>等. 判别式方法将目标跟踪问题看作前景目标与背景的二分类问题, 因此需要产生正、负训练样本进行分类器的构造和更新, 如: 利用 Boosting 或者 Adaboost 的跟踪方法<sup>[9-10]</sup>, 利用支持向量机的跟踪方法<sup>[11-12]</sup>和基于贝叶斯分类的跟踪方法<sup>[13-14]</sup>等.

近年来, 基于稀疏表示的产生式跟踪方法得到了广泛的关注, 并取得了较好的跟踪效果. Mei 等<sup>[8]</sup>提出一种  $L_1$  跟踪器, 将稀疏表示技术引入到运动目标跟踪任务当中, 将跟踪问题视为粒子滤波下的稀疏编码问题, 该方法能够通过微小模板的稀疏表示来解决目标的形变、遮挡和光照变化等问题, 并根据重建误差大小确定最终的跟踪结果. 随后人们又提出了一些改进方法, Zhang 等<sup>[15]</sup>将目标模板线性表示为

**收稿日期:** 2014-07-28; **修回日期:** 2014-11-27.

**基金项目:** 国家自然科学基金项目(61273237, 61271121, 61403116); 中国博士后基金项目(2014M560507); 中央高校基本科研业务费专项资金项目(2013HGBH0045).

**作者简介:** 张旭东(1966—), 男, 教授, 从事机器视觉、传感器技术、智能信息处理及其相关应用系统等研究; 陈仲海(1987—), 男, 硕士生, 从事运动目标跟踪的研究.

所有候选目标的线性组合, 并通过稀疏系数确定目标位置; 针对原微小模板仅适用于部分遮挡的状况, Zhang等<sup>[16]</sup>提出了改进的微小模板; 针对跟踪的效率问题, Liu等<sup>[17]</sup>提出了利用在线特征选择来降低特征的维数; 针对复杂场景的目标跟踪, 潘晴等<sup>[18]</sup>提出了一种构造背景字典的改进方法. 在这些基于稀疏表示的跟踪方法中, 由于只利用原始目标模板来对粒子稀疏表示, 当场景中含有与目标模板相似的背景时容易出现跟踪的漂移.

本文在文献[8]基于稀疏表示跟踪方法的基础上, 基于目标的局部LBP特征, 从当前帧粒子中随机选取一定数目的粒子, 生成粒子模板, 构建一个结合原始目标模板与粒子模板的联合目标函数, 通过对每个生成的粒子进行迭代计算, 并求出稀疏系数; 最后通过粒子加权得到当前帧的目标跟踪位置, 并在跟踪过程中, 采用一定的策略对原始目标模板进行更新.

## 1 基于联合模板稀疏表示的跟踪方法

文献[8]采用一个仿射变换图像对跟踪目标建模, 状态变量 $x_t$ 由仿射变换的6个参数表示, 即 $x_t = (m, n, \theta, s, \alpha, \beta)$ , 6个参数分别表示水平方向平移量、垂直方向平移量、旋转、缩放、纵横比以及斜度. 利用仿射变换, 可以由参数 $x_t$ 从图像中获得感兴趣区域 $z_t$ , 并将 $z_t$ 规范化为与原始目标模板同样的尺寸. 使用高斯分布模拟状态变化分布 $p(x_t|x_{t-1})$ , 并假设仿射变换的6个参数之间是相互无关的. 观测模型 $p(z_t|x_t)$ 反映粒子与模板目标之间的相似度, 通过求解原始目标模板的稀疏表示获得.

该方法中, 原始目标模板集 $T = [t_1, t_2, \dots, t_n] \in R_n^d (d \gg n)$ 表示 $n$ 个原始目标模板, 每个模板 $t_i \in R^d$ , 跟踪结果 $y \in R^d$ , 可以表示为

$$y = Ta + \varepsilon = [T, I] \begin{bmatrix} a \\ e \end{bmatrix}. \quad (1)$$

其中:  $a = [a_1, a_2, \dots, a_n]^T \in R^n$ 为目标稀疏系数,  $\varepsilon$ 中的一部分非零分量表示 $y$ 中的像素被遮挡或者噪声污染, 通过使用微小模板 $I$ 来确定遮挡的位置. 式(1)的稀疏解可以通过求解 $l_1$ 最小化问题得到, 即

$$\min \|Bc - y\|_2^2 + \lambda \|e\|_1. \quad (2)$$

其中:  $B = [T, I]$ ,  $c = [a^T, e^T]^T$ ,  $B$ 为构造的超完备字典, 包含正常模板和遮挡模板. 通过内点法求解 $l_1$ 最小化问题, 投影到原始目标模板空间, 根据 $\|y - Ta\|_2$ 找到最小残差, 进而得到跟踪结果, 跟踪结果即为包含最大概率的状态样本.

当场景中含有与目标相似的背景时, 该方法容易出现跟踪漂移的问题, 如图1所示.  $T_t$ 为图像局部归一化LBP特征向量构成的原始目标模板,  $D_t$ 为当前帧已生成的部分粒子经仿射、局部归一化LBP特征

向量构成的粒子模板. 虚线框为当前帧目标的真实位置, 其他框为当前帧的随机粒子, 其中最右边的实线框对应的粒子与模板中的一个模板非常相似, 在确定目标跟踪位置时, 该粒子可能具有最小的残差, 从而造成跟踪的漂移.

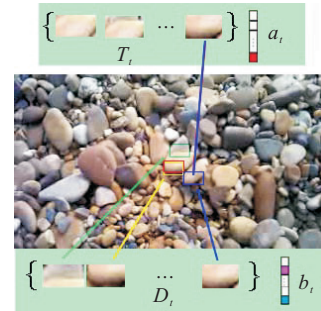


图1 稀疏表示模型

本文将目标外观模型同时用原始目标模板与当前帧部分粒子构成的联合模板稀疏表示, 利用局部归一化LBP特征替代微小模板解决遮挡和光照等问题. 通过迭代求出稀疏系数, 然后对粒子模板加权求和得到最优的状态向量, 解决利用具有最小残差的粒子作为跟踪结果的不稳定性, 在一定程度上解决背景相似产生的跟踪漂移问题.

本文构建一个结合原始目标模板与粒子模板的联合目标函数

$$J(y_t, a_t, b_t) = (\|y_t - T_t a_t\|_2^2 + \lambda_1 \|a_t\|_1) + (\bar{b}_{t-1}/\bar{a}_{t-1})(\|y_t - D_t b_t\|_2^2 + \lambda_2 \|b_t\|_1), \quad (3)$$

其中 $\bar{a}_{t-1}$ 和 $\bar{b}_{t-1}$ 分别为前一帧原始目标模板与粒子模板稀疏系数的最大值. 设定目标模板的权重为1, 若 $\bar{b}_{t-1}/\bar{a}_{t-1} > 1$ , 则表明粒子模板中有比目标模板更接近真实目标的候选粒子, 在下一帧时, 对目标函数中粒子模板赋予更大权重, 否则, 对目标模板赋予更大的权重. 跟踪过程可以看作完成目标函数的最小化

$$\begin{aligned} \{y_t, a_t, b_t\} &= \arg \min_{\{y_t, a_t, b_t\}} J(y_t, a_t, b_t), \\ \text{s.t. } a_t &\geq 0, b_t \geq 0. \end{aligned} \quad (4)$$

由于式(4)中 $y_t$ 、 $a_t$ 和 $b_t$ 均未知, 是一个欠定方程, 本文利用 $y_t = y_{t-1}$ 作为初始条件, 通过一个迭代过程求解式(4)的最优解. 记 $\mu' = \bar{b}_{t-1}/\bar{a}_{t-1}$ , 迭代过程如下.

**Step 1:** 固定 $y_t$ , 得到 $a_t$ 和 $b_t$ . 如果 $y_t$ 已知, 则式(4)的问题可以分成两个标准的 $l_1$ 正则化问题, 利用文献[19]的方法获得最优系数 $a_t$ 和 $b_t$ .

**Step 2:** 固定 $a_t$ 和 $b_t$ , 得到 $y_t$ . 如果参数 $a_t$ 和 $b_t$ 已知, 则式(4)的最优化问题可以重写为

$$y_t = \arg \min_{y_t} \|y_t - T_t a_t\|_2^2 + \mu' \|y_t - D_t b_t\|_2^2. \quad (5)$$

式(5)是一个标准的最小二乘法问题,可以通过下式获得其封闭解:

$$y_t = (T_t a_t + \mu' D_t b_t) / (1 + \mu'). \quad (6)$$

通过以上步骤,式(4)可以通过 Step 1 和 Step 2 的迭代过程求解. 迭代过程的终止条件是  $\|J_t^i - J_t^{i-1}\|_2 \leq \varepsilon$ , 或者迭代到预先设置的迭代次数. 对于跟踪的序列图像, 本文方法平均不超过 8 次迭代即可达到收敛. 最后, 通过带有稀疏系数权重的粒子, 获得目标的跟踪位置

$$x_t = \sum_{i=1}^m \beta_t^i x_t^i, \quad (7)$$

其中  $\beta_t^i = b_t^i / \sum_{i=1}^m b_t^i$ , 即每个粒子的权重为原始权重归一化后的值.

虽然连续两帧之间目标的外观变化不大, 但随着跟踪的连续, 目标的形态、所处的光照等内外因素的影响可能使外观发生较大的改变, 使得初始设定的  $n$  个原始目标模板逐渐不能捕获到目标外观的变化而造成跟踪的逐渐漂移. 为减少跟踪过程中固定的原始目标模板可能带来跟踪偏差的问题, 需要在跟踪的过程中根据当前帧跟踪的结果对原始目标模板更新. 本文利用类似于文献 [8] 的模板更新方法, 在每次更新时, 将目标当前状态的外观相似度与给定的阈值进行比较, 如果小于一定阈值, 则认为原始模板需要更新, 将跟踪结果替代模板中稀疏系数最小的模板.

本文算法的主要步骤如下.

**Step 1:** 原始目标模板初始化. 在第 1 帧图像中, 手工确定被跟踪目标的 3 个坐标点(左上、左下和右上), 通过对手工选定的初始目标模板向 4 个方向移动若干个像素, 产生  $n$  个模板坐标, 形成目标原始模板空间. 将模板集中的  $n$  个图像转换成规定的大小, 应用 LBP 算子获得原始模板集的局部归一化 LBP 特征, 最后将模板集中的 LBP 特征转换为列向量  $T$ .

**Step 2:** 随机粒子生成. 利用仿射变换矩阵, 按高斯分布生成  $N$  个粒子, 得到  $N$  个粒子仿射变换参数.

**Step 3:** 从粒子中随机选取  $m$  个粒子, 生成规定大小的图像, 应用 LBP 算子生成各粒子的局部归一化 LBP 特征, 将这些粒子的 LBP 特征转换为列向量  $D$ .

**Step 4:** 利用本文的迭代算法得到每个粒子对应于稀疏表示的系数  $a_t$  和  $b_t$ , 并利用式(7)得到当前帧目标的跟踪位置.

**Step 5:** 显示当前帧的目标跟踪结果, 对模板进行更新后返回 Step 2.

## 2 实验分析

为了验证本文所提出跟踪方法的有效性, 对一些标准的视频数据集进行跟踪测试. 算法在 Matlab

2009a 下编程实现, 原始目标模板空间的模板数  $n = 10$ , 生成初始目标模板集的扰动像素大小取 3, 粒子滤波器中的粒子取 300, 粒子模板数  $m = 50$ , 联合目标函数中  $\lambda_1 = \lambda_2 = 0.05$ ,  $\varepsilon = 0.01$ , 最大迭代次数设置为 15. 对于每个视频序列, 初始帧的目标位置根据真实值手工标记.

本文算法对 8 个公开的视频图像序列进行了跟踪实验, 分别为具有严重遮挡的 Occlusion1 和 Girl 序列、目标快速运动的 Deer 和 Jumping 序列、具有严重光照变化的 DavidIndoor 和 Singer1 序列、具有相似背景的 DavidOutdoor 和 Stone 序列, 并与跟踪方法中的部分优秀算法 IVT<sup>[4]</sup>、 $L_1$ <sup>[8]</sup>、VTD<sup>[7]</sup>、MIL<sup>[9]</sup>与 CT<sup>[13]</sup>进行了比较, 结果如图 2 所示.

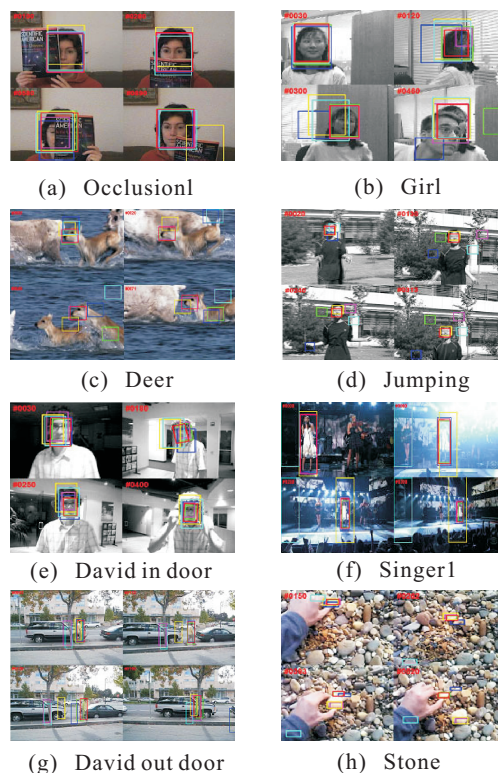


图 2 不同序列的跟踪结果

1) 严重遮挡. 图 2(a) 和图 2(b) 分别为严重遮挡的 Occlusion1 和 Girl 序列跟踪结果. 在 Occlusion1 序列中, 当序列第 890 帧遮挡物书本从人的面前移走后, 除了 MIL 跟踪窗漂移到了书本上, 其他算法整体能在整个序列中完成对目标跟踪, 但本文算法跟踪效果比其他算法都要好; 在 Girl 序列, 包含遮挡、尺度变化和相似目标等难题, 由于 MIL 使用 Haar-like 特征对目标描述, 当目标被相似物遮挡时跟踪效果较差, 而  $L_1$  跟踪方法在目标被相似物遮挡时, 也渐渐漂远离了真实目标.

2) 快速运动. 图 2(c) 和图 2(d) 分别为快速运动的 Deer 和 Jumping 序列跟踪结果. 这两个序列的目标除了运动快速外, 还伴有运动模糊的问题. 在 Deer 序列

中, 本文方法和 VTD 均有较好的跟踪效果, 本文方法略好于 VTD; 在 *Jumping* 序列中, 本文方法要好于其他跟踪方法, MIL 也能完成目标的跟踪, 但效果差一些, 其他 4 种方法的跟踪窗口则逐渐漂移到背景上.

3) 光照变化. 图 2(e) 和图 2(f) 分别为光照变化的 *DavidIndoor* 和 *Singer1* 序列跟踪结果. 这两个序列的目标除了光照变化很大以外, 还伴有目标尺度、姿态的变化. 这两个序列中, 本文方法和 IVT 方法均取得较好的跟踪效果, 其他方法(除 *Singer* 序列中 CT 方法外) 虽然没有出现漂移, 但是准确度不如本文算法.

4) 相似背景. 图 2(g) 和图 2(h) 分别为背景相似的 *David out door* 和 *Stone* 序列跟踪结果, 场景内都含有与目标相似的背景. 在 *David out door* 序列中, VTD、CT、MIL 和  $L_1$  跟踪窗口分别在 93、93、112、128 帧漂

移到不同的背景上, 在第 190 帧时, IVT 跟踪窗口也漂移到背景上, 而本文方法能够始终准确地对目标进行跟踪. 在 *Stone* 序列中, 操作者将目标移到第 150 帧的位置后, CT 方法便漂移到了操作者的手上, 在第 262 帧时, 本文方法和 IVT 方法具有较好的跟踪效果,  $L_1$  跟踪窗口由于含有与目标相似的背景而漂移到了背景石头上, 当将一块背景石头移到目标附近后离开(第 393 帧), VTD 和 MIL 跟踪窗口则漂移到了该背景石头上.

为了定量地比较本文跟踪算法与其他跟踪算法的效果, 计算这 6 种跟踪方法在各个视频序列中每一帧的跟踪误差(即算法跟踪到的某帧的目标中心位置与该帧目标的真实中心位置之间的偏差)和平均跟踪误差, 如图 3 和表 1 所示.

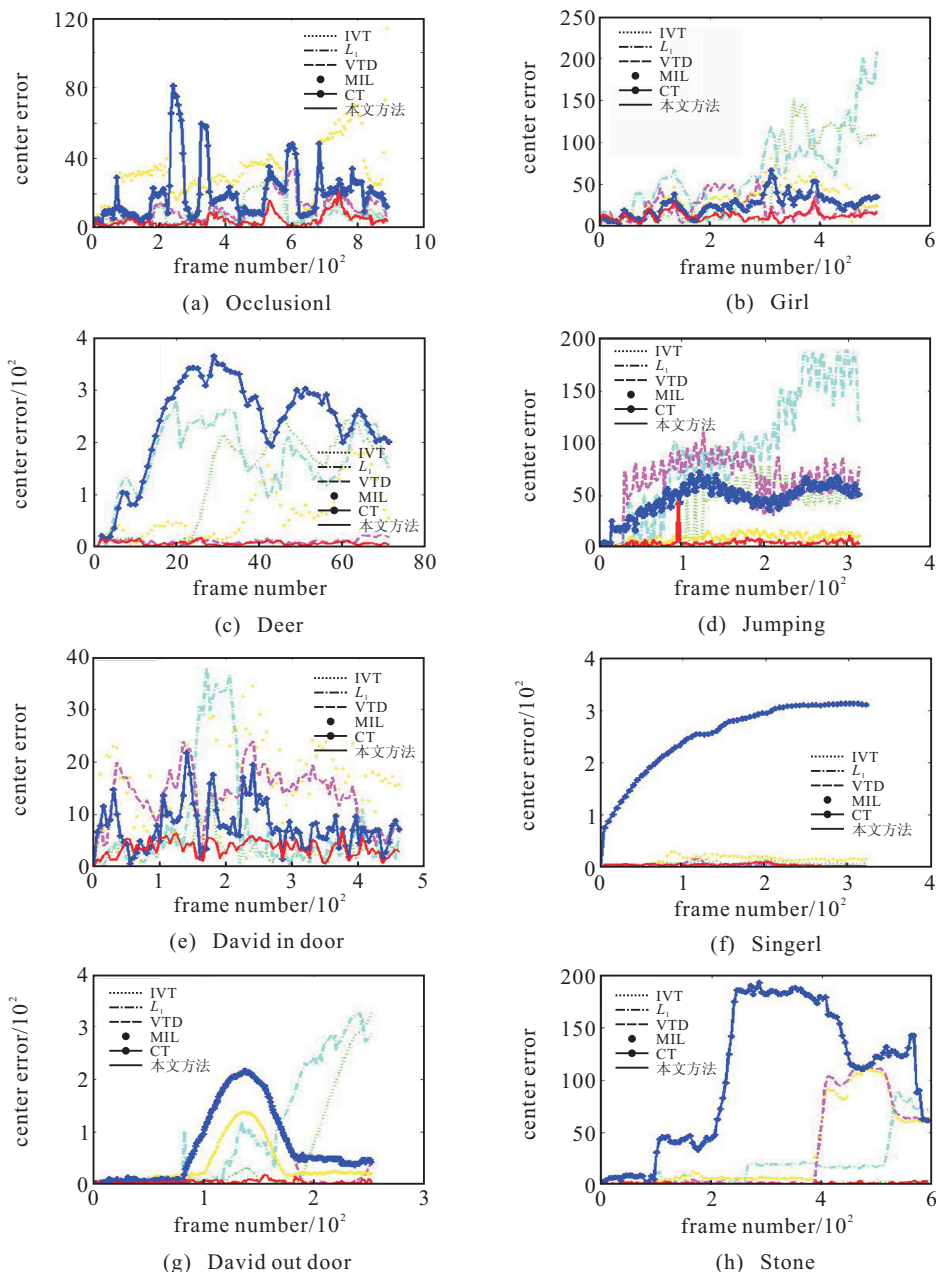


图 3 中心误差对比

表1 平均跟踪误差

	IVT	$L_1$	VTD	MIL	CT	本文方法
Occlusion1	9.2	6.5	11.1	32.3	17.6	<b>5.0</b>
Girl	48.5	62.4	21.4	32.2	25.2	<b>12.9</b>
Deer	127.5	171.5	11.9	66.5	227.1	<b>8.9</b>
Jumping	36.8	92.4	63.0	9.9	47.0	<b>5.2</b>
David in door	<b>3.6</b>	7.6	13.6	16.1	7.8	3.9
Singer1	8.5	4.6	<b>4.1</b>	15.2	250.6	4.8
David out door	53.0	100.3	62.0	38.4	73.0	<b>6.1</b>
Stone	2.3	19.2	31.4	32.3	104.1	<b>1.9</b>
平均值	36.2	58.0	27.3	30.4	94.1	<b>6.1</b>

由图3和表1可见,本文算法的整体跟踪效果远好于其他跟踪方法,能够在目标出现遮挡、快速运动、光照变化和相似背景的情况下较为准确地完成对目标的跟踪,而 $L_1$ 算法因其稳定性较差,跟踪效果整体较差,平均跟踪误差仅小于CT跟踪方法。

### 3 结论

本文提出了一种改进的基于稀疏表示的目标跟踪方法,基于LBP纹理特征,将目标外观模型同时用原始目标模板与当前帧部分粒子构成的联合模板稀疏表示,构建一个结合原始目标模板与粒子模板的联合目标函数,这样,跟踪问题便成为通过迭代求解最优化问题,最后通过带有稀疏系数权重的粒子完成对目标的跟踪。实验结果表明,本文算法能够在序列中存在严重遮挡、光照变化、背景与目标相似等情况下,实现对刚体或非刚体的鲁棒跟踪。但本文算法在跟踪过程中将部分粒子对应的图像作为模板,增加了运算的复杂度,如果使用的粒子增多,虽然效果比较好,但实时性变差。下一步将对如何在保证跟踪准确度的同时提高算法速度进行研究。

### 参考文献(References)

- [1] Yilmaz A, Javad O, Shah M. Object tracking: A survey[J]. ACM Computer Survey, 2006, 38(4): 1-45.
- [2] Adam A, Rivlin E, Shimshoni I. Robust fragments-based tracking using the integral histogram[C]. Proc of IEEE Conf on Computer Vision and Pattern Recognition. New York, 2006: 798-805.
- [3] 杨欣, 费树岷, 李刚, 等. 基于复杂特征融合的改进 mean shift 目标跟踪[J]. 控制与决策, 2014, 29(7): 1297-1300. (Yang X, Fei S M, Li G, et al. Improved mean shift tracking algorithm based on complicated feature fusion[J]. Control and Decision, 2014, 29(7): 1297-1300.)
- [4] Ross D, Lim J, Lin R S, et al. Incremental learning for robust visual tracking[J]. Int J of Computer Vis, 2008, 77(1/2/3): 125-141.
- [5] Wang D, Lu H C, Chen Y W. Incremental MPCA for color object tracking[C]. Int Conf on Pattern Recognition. Istanbul, 2010: 1751-1754.
- [6] Hu W M, Li X, Zhang X, et al. Incremental tensor subspace learning and its applications to foreground segmentation and tracking[J]. Int J of Computer Vision, 2011, 91(3): 303-327.
- [7] Kwon J, Lee K M. Visual tracking decomposition[C]. Proc of IEEE Conf on Computer Vision and Pattern Recognition. San Francisco, 2010: 1269-1276.
- [8] Mei X, Ling H B. Robust visual tracking using  $L_1$  minimization[C]. Proc of IEEE Conf on Computer Vision. Kyoto, 2009: 1436-1443.
- [9] Babenko B, Yang M H, Belongie S. Visual tracking with online multiple instance learning[C]. Proc of IEEE Conf on Computer Vision and Pattern Recognition. Anchorage, 2009: 983-990.
- [10] 徐建军, 张蓉, 毕笃彦, 等. 一种新的 AdaBoost 视频跟踪算法[J]. 控制与决策, 2012, 27(5): 681-685. (Xu J J, Zhang R, Bi D Y, et al. A new AdaBoost video tracking algorithm[J]. Control and Decision, 2012, 27(5): 681-685.)
- [11] Avidan S. Support vector tracking[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2004, 26(8): 1064-1072.
- [12] Fu K, Gong C, Qiao Y, et al. One-class support vector machine-assisted robust tracking[J]. J of Electronic Imaging, 2013, 22(2): 023002.
- [13] Zhang K H, Zhang L, Yang M S. Real-time compressive tracking[C]. Proc of IEEE Conf on European Computer Vision. Florence, 2012: 866-879.
- [14] 罗会兰, 钟宝康, 孔繁胜. 结合目标预测位置的压缩跟踪[J]. 中国图象图形学报, 2014, 19(6): 857-885. (Luo H L, Zhong B K, Kong F S. Object tracking algorithm by combining the predicted target position with compressive tracking[J]. J of Image and Graphics, 2014, 19(6): 857-885.)
- [15] Zhang S P, Yao H X, Sun X, et al. Robust object tracking based on sparse representation[C]. Proc of SPIE Int Conf on Visual Communications and Images Processing. Huangshan, 2010: 77441N-1-8.
- [16] Zhang S P, Yao H X, Zhou H Y, et al. Robust visual tracking based on online learning sparse representation[J]. Neurocomputing, 2013(100): 31-40.
- [17] Liu B Y, Yang L, Huang J Z. Robust and fast collaborative tracking with two stage sparse optimization[C]. Proc of IEEE Conf on European Computer Vision. Heraklion, 2010: 624-637.
- [18] 潘晴, 曾仲杰. 基于稀疏表示和特征选择的LK目标跟踪[J]. 计算机应用研究, 2014, 31(2): 623-628. (Pan Q, Zeng Z J. LK tracking based on sparse representation and features selection[J]. Application Research of Computers, 2014, 31(2): 623-628.)
- [19] Tibshirani R. Regression shrinkage and selection via the lasso[J]. J of the Royal Statistical Society, 1996: 267-288.