

DOI:10.19322/j.cnki.issn.1006-4710.2017.03.002

# 基于多模态词典学习的目标跟踪算法

王 婧, 朱 虹

(西安理工大学 自动化与信息工程学院, 陕西 西安 710048)

**摘要:** 稀疏表示在目标跟踪中已取得了良好的跟踪效果,但过完备词典模式单一、数据庞大,稀疏系数需用复杂的优化算法求解,会限制此类算法的跟踪性能提高。因此,本文在粒子滤波框架下,提出了一种基于多模态词典的目标跟踪算法。首先,创建长、短周期的正样本模板,结合负样本模板共同构成多模态词典,用以表征采样目标当前状态;其次,根据样本与词典之间的多模态相关系数,对目标进行粗跟踪,得到候选跟踪结果;最后,利用 Local Maximal Occurrence(LOMO)特征构建候选跟踪目标与多模态词典的观测似然函数,取具有最大似然度的候选跟踪目标作为最终的跟踪结果。实验结果表明,本文算法在遮挡、光照变化和背景干扰情况下均具有较强的跟踪鲁棒性。

**关键词:** 目标跟踪;多模态词典;粒子滤波;稀疏表示

**中图分类号:** TP391.4

**文献标志码:** A

**文章编号:** 1006-4710(2017)03-0259-06

## Object tracking based on multi-modality dictionary learning

WANG Jing, ZHU Hong

(School of Automation and Information Engineering, Xi'an University of Technology, Xi'an 710048, China)

**Abstract:** Sparse representation-based methods have been successfully applied to visual tracking. However, the over complete dictionary mode is single and data is large, and the sparse coefficients need to be solved by the complex optimization algorithm, which will limit their tracking performances. In this paper, within the tracking framework of particle filter, we propose a tracking method based on the multi-modality dictionary learning. Firstly, a long and short period of object templates are created, combined with background templates to form a multi-modality dictionary to characterize the current state of sampled object. Secondly, according to the multi-modal coefficients between the sampled objects and the dictionary, the target is tracked roughly with the candidate tracking results obtained. Finally, the observation likelihood functions of the candidate tracking results and the multi-modality dictionary are constructed by using LOMO features, and the candidate tracking result with the maximum likelihood is taken as the final tracking result. Experimental results demonstrate that the proposed method has strong tracking robustness in the case of occlusion, illumination change and background interference.

**Key words:** object tracking; multi-modality dictionary; particle filter; sparse representation

目标跟踪是计算机视觉领域中重要的一个分支,其已经在很多领域得到了广泛的应用,比如视频监控、智能交通管控、医学图像和军事模拟等<sup>[1]</sup>。虽然目标跟踪的发展历史已有几十年,并取得了许多有目共睹的成绩,但由于跟踪结果易受到目标姿态和尺度变化的影响,以及背景复杂、光照变化等外界因素的干扰,造成跟踪结果不理想。所以,复杂状况

下的高精度目标跟踪依然是具有挑战性的研究课题。

目标跟踪方法大致可分为:生成式跟踪算法和判别式跟踪算法两类<sup>[2]</sup>。生成式方法是将与外观模型具有最佳相容性的候选对象作为跟踪目标<sup>[3-5]</sup>。判别式跟踪算法将目标跟踪视为将前景目标与背景进行二分类的问题<sup>[6-8]</sup>。一般来说,前者可以获得更

**收稿日期:** 2016-07-12

**基金项目:** 国家自然科学基金资助项目(61673318,61771386);陕西省自然科学基金基础研究计划资助项目(2016JM6045);陕西省教育厅科学研究计划专项资助项目(16JK1571)

**作者简介:** 王婧,女,博士生,研究方向为视频图像处理、模式识别。E-mail: jing8563@163.com

**通讯作者:** 朱虹,女,教授,博士,研究方向为数字图像处理、智能视频监控、模式识别。E-mail: zhuhong@xaut.edu.cn

精确的目标特征,但计算复杂度较高。后者可以获得更好的跟踪精度,但必须预先处理大量的训练样本,一旦外观发生了改变,就需要重新训练样本,训练样本不足极易造成跟踪失败。

另外,随着机器学习理论的不断深入,目前跟踪准确度较高的算法大多是基于深度卷积网络<sup>[9-11]</sup>。深度卷积网络需要用非常庞大的训练数据,以保证目标特征的有效性。这对于目标跟踪等视频应用来说,收集这些数据是比较困难的。另一方面,大量的卷积运算也会使跟踪速度变慢,降低跟踪的实时性。

近年来,稀疏表示类跟踪方法表现出了良好的跟踪能力<sup>[12-16]</sup>。该算法一般通过建立过完备词典作为特征空间来描述目标外观。当选择了合适的词典时,目标的高维数据可以稀疏表示<sup>[13,16-18]</sup>。但是这类方法本质上是  $\ell_1$  最小化问题,其计算量较大,并且需要复杂的优化算法计算出准确的稀疏编码,此外目标间的相互关系也容易被忽略。

本文提出基于多模态词典的目标跟踪算法,主要创新点为:构建多模态词典,在跟踪过程中加入了目标间的相关信息,能够准确地表示目标,省去了构建过完备词典的冗余过程;将目标跟踪问题简化为词典学习问题,通过简单计算目标与词典间的系数,实现目标粗跟踪;将用于行人再识别的 LOMO 特征应用在目标跟踪中,在不采用深度卷积网络的情况下,实现目标的精确跟踪。

## 1 粒子滤波

本文使用粒子滤波作为跟踪框架,根据每一帧中粒子的观测概率去估计下一帧的目标<sup>[19]</sup>。假设  $Y_t = [y_1, \dots, y_t]$  是第一帧到第  $t$  帧的观测图像,  $x_t$  是用于描述时刻  $t$  的目标运动参数的状态变量,则  $x_t$  的粒子滤波遵循如下的概率分布:

$$p(x_t | Y_t) \propto p(y_t | x_t) \int p(x_t | x_{t-1}) p(x_{t-1} | Y_{t-1}) dx_{t-1} \quad (1)$$

其中,  $p(x_t | x_{t-1})$  为状态转移分布,  $p(y_t | x_t)$  用来估计观测值  $y_t$  在状态  $x_t$  时的似然函数。

采样粒子以上一帧跟踪结果的中心坐标作为中心,按高斯函数分布进行采样。由于粒子的数量会影响跟踪效率,所以一般情况下都需要过滤无关采样粒子,降低跟踪冗余度。

## 2 基于多模态词典的目标跟踪算法

### 2.1 目标描述

Liao 等人针对行人再识别中目标在不同相机

中的特征表现不一致,提出了 Local Maximal Occurrence(LOMO)特征<sup>[20]</sup>,这是近年来可以与深度学习网络特征相媲美的一种有效的特征。由于跟踪过程中的目标与行人再识别中的目标所面临的问题极为相似,鉴于 LOMO 特征的有效性已经验证,因此,本文使用 LOMO 来描述目标特征,用于表征目标的观测似然度函数。

LOMO 特征首先对目标的彩色图像使用多尺度 Retinex 算法,使得不同场景下的目标具有颜色一致性,然后对其提取 HSV 颜色直方图作为目标的颜色特征。随后,使用 Scale Invariant Local Ternary Pattern(SILTP)描述子<sup>[21]</sup>可以得到更具鲁棒性的目标特征。

### 2.2 多模态词典的构建与更新

#### 2.2.1 多模态词典的构建

稀疏表示类跟踪方法会使用过完备词典对目标进行稀疏编码。稀疏编码过程包含了两个问题:一是稀疏编码,用于求解可以表示目标与已知词典关系的系数;二是词典学习,用于解决词典构造的问题<sup>[22]</sup>。基于稀疏表示原理,候选目标  $x_i$  可以表示为对应词典  $D$  的稀疏编码  $\alpha_i$  的线性组合。 $x_i$  对应的稀疏编码  $\alpha_i \in \mathbf{R}^{n+m}$  满足以下条件:

$$\min_{\alpha_i} \frac{1}{2} \|x_i - D\alpha_i\|_2^2 + \lambda \|\alpha_i\|_1 \quad (2)$$

其中,过完备词典  $D = [D_p, D_n] \in \mathbf{R}^{d \times (n+m)}$  由前景词典  $D_p \in \mathbf{R}^{d \times n}$  和背景词典  $D_n \in \mathbf{R}^{d \times m}$  构成。该问题属于 Lasso 回归<sup>[23]</sup>问题,候选目标  $x_i$  的稀疏编码  $\alpha_i$  一般可以由 LARS 方法<sup>[24]</sup>计算得到。

然而,这种求解方法不仅需要数量庞大的词典,而且更新优化过程比较复杂。事实上,对于没有受到严重外界影响的正常状态下的目标,仅需要少量词典数量即可将目标从背景中区分开。而当目标受到外界环境影响或者自身姿态发生变化时,仅依靠相邻帧的图像信息去寻找目标也是远远不够的。所以,如果在跟踪过程中可以实时得到适合表征当前目标的词典,就无需冗余的过完备词典的构建以及复杂的迭代过程。

本文中,正样本词典由短周期目标模板  $D_s$  和长周期目标模板  $D_l$  构成,负样本词典是在视频图像中随机选取的非目标区域  $D_n$ 。因此,多模态词典  $D$  由两部分构成,即  $D = [D^s, D^l]$ ,其中  $D^s = [D_s, D_n] \in \mathbf{R}^{d \times (m_s + n_s)}$ ,  $D^l = [D_l, D_n] \in \mathbf{R}^{d \times (m_l + n_l)}$ 。词典模板由观测目标的像素值表示。 $D_s$  由已知的第一帧目标位置来初始化,分别向上、下、左、右四个方向

移动若干个像素。 $D_l$  为  $D_s$  的聚类中心,即  $D_l = \frac{1}{l_s} \sum_{i=1}^{l_s} D_s^{(i)}$ ,  $l_s$  表示  $D_s$  中样本模板的个数。需要注意词典中每个模板的观测向量是归一化的。

### 2.2.2 多模态词典的更新

构建多模态词典,一方面在目标外观变化不大时,短周期词典可以有效将目标从背景中区分开,长周期词典可以防止误差的累积;另一方面在目标外观变化较大时,利用短周期词典可以连续跟踪目标,使用长周期词典防止正确的采样目标丢失,两种模式的词典相结合可以更好地平衡  $\ell_1$  跟踪算法中的适应性和鲁棒性。因此,多模态词典的更新就显得很重要。 $D_s$  使用前一帧的目标来进行学习和更新, $D_l$  使用前面所有帧的目标来进行学习和更新。 $D_l$  更新时,首先对当前帧之前的所有帧的正负样本模板进行 K-means 聚类,求解其聚类中心作为长周期词典。按照公式(3),计算当前跟踪目标  $x^t$  与长周期词典模板聚类中心  $D_l$  的欧氏距离来判断当前目标归属于哪一类,之后对所属类别的聚类中心进行更新,获得新的长周期词典。

$$\begin{cases} \{x^t \in D_l^i, d(f(x^t), f(D_l^i)) \in [0, d_{\max} + \Delta]\} \\ \{x^t \in D_l^{\text{new}}, d(f(x^t), f(D_l^{\text{new}})) > d_{\max} + \Delta\} \end{cases} \quad (3)$$

其中, $D_l^i$  表示长周期词典中不同的类别模板, $D_l^{\text{new}}$  表示新的聚类词典, $d(\cdot)$  表示欧氏距离, $f(\cdot)$  表示提取 LOMO 特征, $d_{\max}$  表示初始短周期词典  $D_s$  模板间的欧式距离  $d(D_s, D_s')$  的最大值, $\Delta$  为一个常数变量。

### 2.3 目标跟踪算法

本文跟踪算法采用粒子滤波跟踪框架,所有采样粒子表示为  $Y = \{y_1, y_2, \dots, y_N\} \in \mathbf{R}^{d \times N}$ ,使用距离约束过滤无关采样粒子,即采样目标的中心坐标  $p(x_t^i)$  与上一帧跟踪结果的中心坐标  $p(x_{t-1})$  之间应该满足  $\|p(x_t^i) - p(x_{t-1})\| \leq \max(w, h)$ ,其中  $w$  和  $h$  代表跟踪窗口的宽和高。过滤后的候选采样目标表示为  $X = \{x^i | i \in [1, q]\} \in \mathbf{R}^{d \times q} (q \ll N)$ 。

假设多模态词典可以适应目标的变化而无需更新,则对于词典中的每个模板来说,其与理想跟踪结果之间的代价函数应该满足最小。采样目标关于短周期词典  $D^s$  和长周期词典  $D^l$  的代价函数分别如公式(4)和公式(5)所示。使用 LARS<sup>[24]</sup> 求取对应的稀疏系数。

$$l_s(x^i, D^s) = \min_{\alpha_s^i} \frac{1}{2} \|x^i - D^s \cdot \alpha_s^i\|_2^2 + \lambda \|\alpha_s^i\|_1 \quad (4)$$

$$l_l(x^i, D^l) = \min_{\alpha_l^i} \frac{1}{2} \|x^i - D^l \cdot \alpha_l^i\|_2^2 + \lambda \|\alpha_l^i\|_1 \quad (5)$$

考虑到理想的采样目标与词典中的正样本模板会越接近,与负样本模板越不接近,同理,不理想的采样目标会与负样本模板越接近。故理想目标满足正样本对应的系数与负样本系数的  $\ell_1$ -norm 差异越大。因此为了提高跟踪结果的准确率,本文取满足公式(7)所示条件的最大前  $p$  个样本作为跟踪候选结果  $R$ ,实现目标的粗跟踪,如公式(6)所示。

$$R = [R_s, R_l] = [I_s^i |_\rho, I_l^i |_\rho] \quad (i \in [1, q]) \quad (6)$$

$$\begin{aligned} I_s^i &= \max(\|\alpha_s^{i+}\|_1 - \|\alpha_s^{i-}\|_1) \\ I_l^i &= \max(\|\alpha_l^{i+}\|_1 - \|\alpha_l^{i-}\|_1) \end{aligned} \quad (7)$$

其中, $\alpha_s^{i+}$  和  $\alpha_s^{i-}$  为短周期词典中正样本和负样本对应的系数, $\alpha_l^{i+}$  和  $\alpha_l^{i-}$  为长周期词典中正样本和负样本对应的系数。

最后,为了得到准确的跟踪结果,为每一个候选跟踪结果建立由目标 LOMO 特征表示的观测似然函数,并与多模态词典进行相似度比较,取具有最大相似度的候选跟踪结果作为最终的跟踪结果,实现目标由粗到细的跟踪,如公式(8)和(9)所示。

$$\hat{x} = \operatorname{argmax}(\omega_s \cdot s_s + \omega_l \cdot s_l) \quad (8)$$

$$\begin{aligned} s_s &= \operatorname{sim}(f(R_s^i), f(D_s)) \\ s_l &= \operatorname{sim}(f(R_l^i), f(D_l)) \end{aligned} \quad (9)$$

其中, $\operatorname{sim}(\cdot)$  表示用巴氏距离计算特征的相似度。 $f(\cdot)$  为对应图像区域的 LOMO 特征。 $\omega_s$  和  $\omega_l$  为对应动态词典下候选跟踪结果相似度对应的权重, $\omega_s = s_s / (s_s + s_l)$ ,  $\omega_l = s_l / (s_s + s_l)$ 。

本文算法流程如下。

- 1) 输入: 第  $t$  帧待跟踪图像
- 2) 输出: 第  $t$  帧图像中的跟踪结果  $\hat{x}_t$
- 3) 初始化: 利用第一帧图像信息构建多模态词典跟踪:

for  $t=2$ : end

① 根据前一帧的跟踪结果分布采样粒子,利用距离约束条件过滤无关粒子;

② 使用 LARS 方法得到采样目标关于多模态词典的稀疏系数;

③ 根据每个采样目标的系数,得到候选跟踪结果  $R$ (公式(6));

④ 计算候选跟踪结果中每一个样本的观测似然度(公式(8));

⑤ 取最大相似度的观测值对应的样本作为最终的跟踪结果;



⑥利用  $K$  均值聚类原理更新多模态词典(第 3.2 节)。  
end for

### 3 实验与分析

#### 3.1 评价准则

文中使用误差率和重叠率两个指标来评价算法。误差率  $error$  即为算法得到的目标中心坐标与理想目标中心坐标之间的欧氏距离,其数值越小,代表算法跟踪的位置越准确。

$$error = \sqrt{(x^{ALG} - x^{GT})^2 + (y^{ALG} - y^{GT})^2} \quad (10)$$

其中,  $(x^{ALG}, y^{ALG})$  为算法得到跟踪结果的中心坐

标,  $(x^{GT}, y^{GT})$  是理想跟踪结果的中心坐标。

重叠率  $OR$  为各种算法得到的跟踪窗口与理想跟踪窗口的重叠比例,其数值越大,代表跟踪窗口越符合目标的大小。

$$OR = |G_i \cap T_i| / |G_i \cup T_i| \quad (11)$$

其中,  $G_i$  是理想跟踪结果的跟踪窗口,  $T_i$  是算法得到的跟踪结果的跟踪窗口。

#### 3.2 实验分析

本文使用的测试视频 FaceOcc1、Walking、Fish 和 Coupon 选自于国际标准测试集<sup>[1]</sup>。IVT<sup>[3]</sup>、L1APG<sup>[13]</sup>、TLD<sup>[7]</sup>、SP<sup>[5]</sup> 和 CF2<sup>[25]</sup> 作为本文的对比算法。不同算法的跟踪结果示意图如图 1 所示。

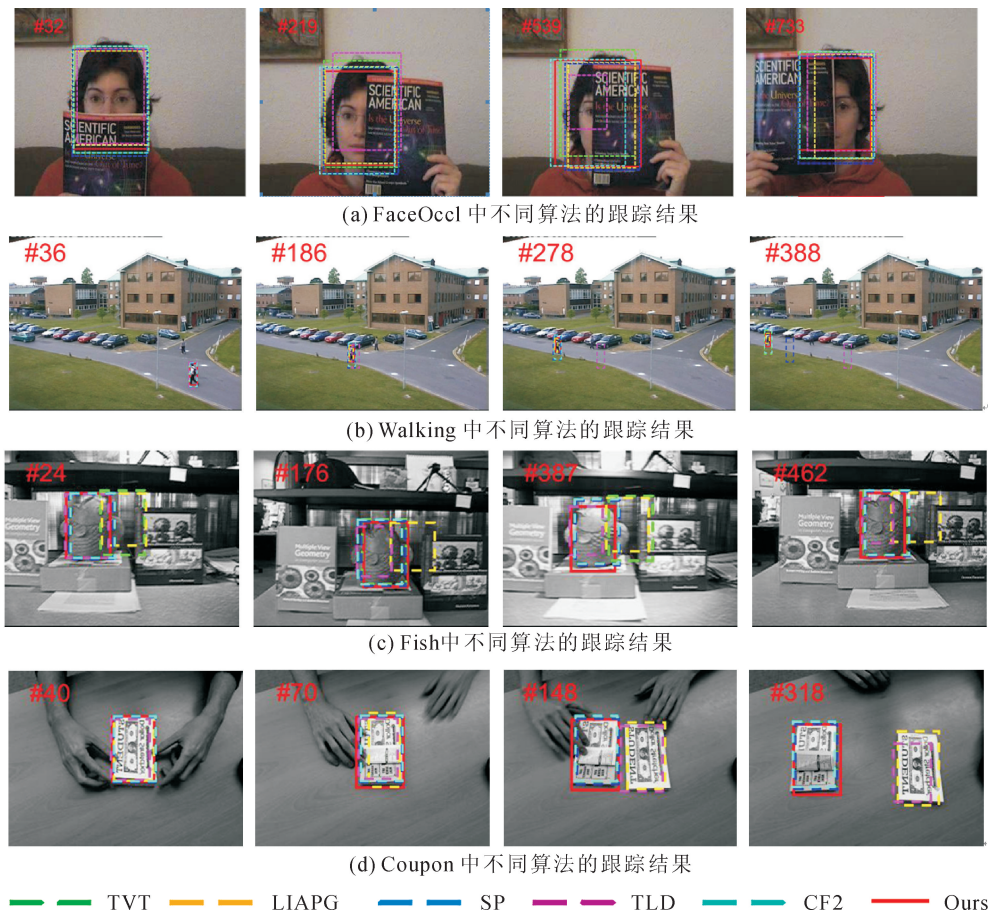


图 1 不同测试视频中各种算法的跟踪结果示意图  
Fig. 1 Some representative results of test sequences

图 2 和图 3 是不同测试数据中各种算法的误差率曲线和重叠率曲线。

测试数据 FaceOcc1 中存在的主要问题是目标受到大面积、长时间的遮挡。TLD 算法在目标受到大面积遮挡时,跟踪失败,但当目标恢复正常状态后能够继续跟踪。IVT、L1APG、SP、CF2 和本文算法的跟踪效果差异不大,均能保持较低的跟踪误差和

较高的重叠率,其中,本文算法效果最好。

测试数据 Walking 中存在的主要问题是局部遮挡和目标尺寸变化。当目标尺寸逐渐变小,目标特征与背景不易区分,TLD 和 SP 丢失目标,IVT、L1APG、CF2 和本文算法一直定位在目标上,其中 IVT 和 CF2 的跟踪窗口过大,不能很好地适应目标大小。L1APG 和本文算法跟踪效果最好。

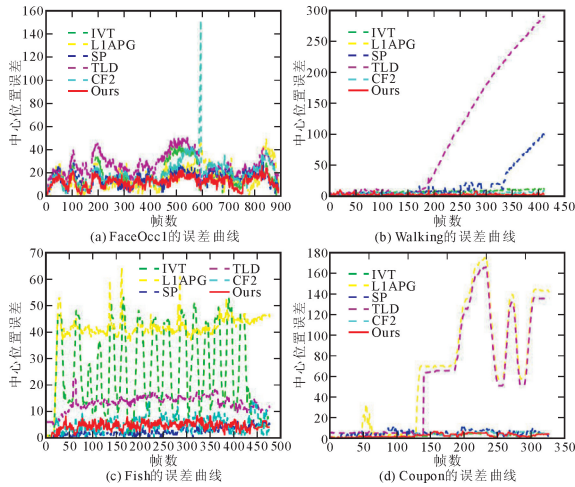


图2 不同测试视频的误差曲线  
Fig. 2 Tracking error plots for all test sequences

测试数据 Fish 中存在的主要问题是光照变化。由于目标不断受到光照和相机视角变化的影响, IVT 和 L1APG 丢失目标。TLD 大致可定位在目标上,但是跟踪窗口过小。SP、CF2 和本文算法跟踪结果大致相同,评价数据上差异不大,表现最优的是 SP,本文算法表现次优。

测试数据 Coupon 中存在的主要问题是相似背景的干扰。当背景中出现与目标外观相似的物体

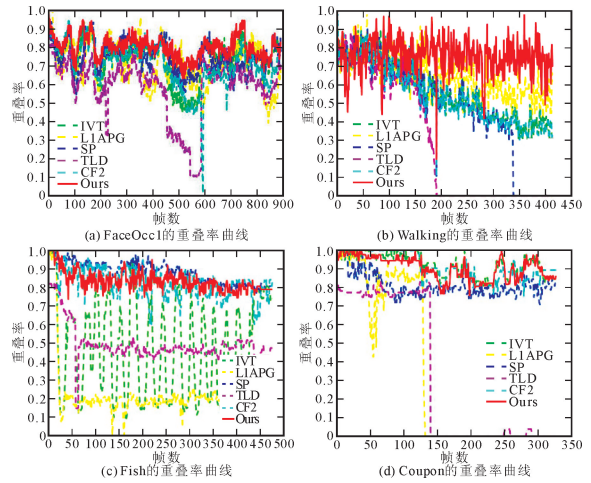


图3 不同测试视频的重叠率  
Fig. 3 Tracking overlap plots for all test sequences

时,L1APG 和 TLD 定位在相似物体上,跟踪失败。IVT、SP、CF2 和本文算法都可以准确的进行目标定位,对应的跟踪效果比较相似。

表 1 是各种算法在不同测试视频中的平均误差值和平均重叠率。从表 1 可看出,本文算法在 FaceOcc1 和 Walking 中保持最优的跟踪效果,在 Fish 和 Coupon 中为次优结果,与最优结果仅存在少量差异。

表 1 各种跟踪算法在测试数据中的平均跟踪误差和平均准确率  
Tab. 1 Mean of tracking error and tracking overlap rate

算法	FaceOcc1		Walking		Fish		Coupon	
	error/像素	OR/%	error/像素	OR/%	error/像素	OR/%	error/像素	OR/%
IVT	17.92	74.92	7.62	56.57	24.42	47.38	<b>2.31</b>	<b>92.05</b>
L1APG	16.99	71.57	<u>2.92</u>	<u>65.34</u>	40.70	21.43	65.12	33.91
SP	<u>14.61</u>	<u>75.13</u>	19.32	48.88	<b>2.48</b>	<b>88.38</b>	4.96	81.39
TLD	27.38	59.69	95.89	29.06	13.22	49.54	59.69	33.26
CF2	19.66	71.73	4.49	55.14	4.88	82.62	2.89	<u>91.65</u>
本文算法	<b>13.73</b>	<b>80.46</b>	<b>2.63</b>	<b>73.29</b>	<u>4.56</u>	<u>83.55</u>	<u>2.69</u>	90.66

说明:表中粗体+下划线数据表示最优结果,斜体+下划线数据表示次优结果。

#### 4 结 语

本文提出了一种基于多模态词典学习的目标跟踪算法。该算法将目标跟踪问题视为词典学习问题。通过构造多模态词典,引入目标间的相互关系,可以更加合理地表征目标,从而目标与词典的多模态系数可以用于寻找候选跟踪结果,再利用 LOMO 特征可以进一步得到准确的跟踪目标。实验结果证明了本文算法无需复杂的优化过程就达到了与卷积网络类跟踪算法(CF2)的跟踪精度,验证了本文构建的多模态词典的有效性,以及在遮挡、光照变化和背景干扰情况下均具有较强的跟踪鲁棒性。

#### 参考文献:

[1] WU Y, LIM J, YANG M H. Online object tracking: a benchmark [C]// IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2013, Portland, USA, 2013: 2411-2418.  
 [2] 尹宏鹏,陈波,柴毅,等. 基于视觉的目标检测与跟踪综述 [J]. 自动化学报, 2016, 42(10): 1466-1489.  
 YIN Hongpeng, CHEN Bo, CHAI Yi, et al. Vision-based object detection and tracking: a review [J]. ACTA Automatica Sinica, 2016, 42(10): 1466-1489.  
 [3] ROSS D A, 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.

- YIN Hongpeng, CHEN Bo, CHAI Yi, et al. Vision-based object detection and tracking: a review[J]. *ACTA Automatica Sinica*, 2016, 42(10):1466-1489.
- [4] WANG J, ZHU H, YU S, et al. Object tracking using color-feature guided network generalization and tailored feature fusion [J]. *Neurocomputing*, 2017, 238: 387-398.
- [5] WANG D, LU H, YANG M H. Online object tracking with sparse prototypes [J]. *IEEE Transactions on Image Processing*, 2013, 22(1): 314-325.
- [6] KALAL Z, MIKOLAJCZYK K, MATAS J. Tracking-learning-detection [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34 (7): 1409-1422.
- [7] KALAL Z, MATAS J, MIKOLAJCZYK K. P-N learning: bootstrapping binary classifiers by structural constraints [J]. *Computer Vision & Pattern Recognition*, 2010, 238(6): 49-56.
- [8] 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.
- [9] WANG L, OUYANG W, WANG X, et al. Visual tracking with fully convolutional networks [C]// *IEEE International Conference on Computer Vision*, December 13-16, 2015, Santiago, Chile, 2015: 3119-3127.
- [10] NAM H, HAN B. Learning multi-domain convolutional neural networks for visual tracking [C]// *IEEE Conference on Computer Vision and Pattern Recognition*, June 27-30, 2016, Las Vegas, USA, 2016: 4293-4302.
- [11] ZHANG K, LIU Q, WU Y, et al. Robust visual tracking via convolutional networks without training [J]. *IEEE Transactions on Image Processing*, 2016, 25(4): 1779-1792.
- [12] ZHONG W, LU H, YANG M H. Robust object tracking via sparsity-based collaborative model [C]// *IEEE Conference on Computer Vision and Pattern Recognition*, June 16-21, 2012, Providence, USA, 2012: 1838-1845.
- [13] WANG N, WANG J, YEUNG D Y. Online robust non-negative dictionary learning for visual tracking [C]// *IEEE International Conference on Computer Vision*, December 1-8, 2013, Sydney, Australia, 2013: 657-664.
- [14] XING J, GAO J, LI B, et al. Robust object tracking with online multi-lifespan dictionary learning [C]// *IEEE International Conference on Computer Vision*, December 1-8, 2013, Sydney, Australia, 2013: 665-672.
- [15] MEI X, LING H. Robust visual tracking and vehicle classification via sparse representation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, 33(11): 2259-2272.
- [16] BAO C, WU Y, LING H, et al. Real time robust L1 tracker using accelerated proximal gradient approach [C]// *IEEE Conference on Computer Vision and Pattern Recognition*, June 16-21, 2012, Providence, USA, 2012: 1830-1837.
- [17] QIAN C, XU Z. Robust visual tracking via sparse representation under subclass discriminant constraint [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2016, 26(7): 1293-1307.
- [18] XIE Y, ZHANG W, LI C, et al. Discriminative object tracking via sparse representation and online dictionary learning [J]. *IEEE Transactions on Cybernetics*, 2014, 44(4): 539-553.
- [19] CHANG C, ANSARI R. Kernel particle filter for visual tracking [J]. *IEEE Signal Processing Letters*, 2015, 12(3): 242-245.
- [20] LIAO S, HU Y, ZHU X, et al. Person re-identification by local maximal occurrence representation and metric learning [C]// *IEEE Conference on Computer Vision and Pattern Recognition*, June 7-12, 2015, Boston, USA, 2015: 2197-2206.
- [21] LIAO S, ZHAO G, KELLOKUMPU V, et al. Modeling pixel process with scale invariant local patterns for background subtraction in complex scenes [C]// *IEEE Conference on Computer Vision and Pattern Recognition*, June 13-18, 2010, San Francisco, USA, 2010: 1301-1306.
- [22] AHARON M, ELAD M, BRUCKSTEIN A. K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation [J]. *IEEE Transactions on Signal Processing*, 2006, 54(11): 4311-4322.
- [23] TIBSHIRANI R. Regression shrinkage and selection via the lasso [J]. *Journal of the Royal Statistical Society*, 2011, 73(3): 273-282.
- [24] TIBSHIRANI R, JOHNSTONE I M, EFRON B, et al. Least angle regression [J]. *Annals of Statistics*, 2004, 32(2): 407-451.
- [25] MA C, HUANG J B, YANG X, et al. Hierarchical convolutional features for visual tracking [C]// *IEEE International Conference on Computer Vision*, December 7-13, 2015, Santiago, Chile, 2015: 3074-3082.

(责任编辑 王卫勋)