

# 一种基于压缩感知的在线学习跟踪算法

刘 威 赵文杰 李 成

空军航空大学航空航天情报系, 吉林 长春 130022

**摘要** 实现稳健的目标跟踪,建立有效的目标在线模型至关重要。针对现有在线学习跟踪算法缺乏对目标观测信息是否有效的判断,提出了一种简单且高效的解决方法。利用正负样本构建目标在线模型,基于压缩感知理论从多尺度图像特征空间提取特征信息完成目标表征之后,由随机森林分类器进行分类并通过一种特征置信度量策略来确定在线更新速率,最后由目标在线模型判断输出置信度最高的样本,此外还建立了一种遮挡反馈机制来决定是否更新目标在线模型。实验结果表明,该方法在目标被长时间遮挡、光照变化等情况下均能完成稳健跟踪,在320 pixel×240 pixel大小的视频序列中处理速度保持在30~50 frame/s左右,可以满足实时应用的需求。

**关键词** 机器视觉; 目标跟踪; 压缩感知; 随机森林分类器; 目标在线模型

**中图分类号** TP391 **文献标识码** A

**doi:** 10.3788/AOS201535.0915001

## An Online Learning Visual Tracking Method Based On Compressive Sensing

Liu Wei Zhao Wenjie Li Cheng

*Department of Aerospace Intelligence, Aviation University of Air Force, Changchun, Jilin 130022, China*

**Abstract** It is crucial to establish an effective online model for robust tracking. As existing online learning tracking algorithms do not judge whether the objective observation information is effective, a simple and efficient solution is proposed. The positive and negative samples are applied to build online object model, then feature information is extracted from the multi-scale image feature space by compressive sensing to represent object, the random fern classifier is adopted to classify and determine the online update rate by a confidence measure strategy of features. The online object model will output the sample with the highest confidence, which is decided whether to update by an shelter feedback mechanism. Experimental results show that the proposed algorithm can complete the robust tracking under the condition of long-time occlusion, illumination changing, the video sequence of 320 pixel×240 pixel, the processing speed can keep 30~50 frame/s, which meets real-time application requirement.

**Key words** machine vision; object tracking; compressive sensing; random fern classifier; online object model

**OCIS codes** 150.4065; 100.4999; 330.4150

## 1 引 言

近年来,大量基于检测的跟踪算法被提出并在特定场景下获得了不错的跟踪效果,但由于遮挡、光照变化、图像模糊而导致目标外观不断变化,使得稳健性的目标跟踪依然是一个很有挑战性的难题<sup>[1]</sup>。一个有效的目标在线模型对稳健性跟踪至关重要,因此引起了学者的广泛关注。按照目标在线模型,目标跟踪可以划分为基于生成式的跟踪算法和基于判别式的跟踪算法。

基于生成式的跟踪算法<sup>[2-4]</sup>通常只利用目标本身信息来构建在线模型,然后基于最小重构误差在图像区域内搜索目标;基于判别式的跟踪算法<sup>[5-7]</sup>将目标跟踪看成一个二元分类问题,其任务是找到将目标从背景

收稿日期: 2015-01-26; 收到修改稿日期: 2015-04-28

基金项目: 国家自然科学基金(61301233)

作者简介: 刘 威(1991—),男,硕士研究生,主要从事视觉跟踪和运动分析方面的研究。E-mail: 1224337250@qq.com

导师简介: 赵文杰(1968—),女,博士,教授,主要从事光学成像自动目标识别方面的研究。E-mail: actionlw@163.com

中分离出来的最优决策边界。

为了充分利用上述两类算法的各自优势,文献[8]基于压缩感知理论提出了一种有效的实时跟踪算法(CT),通过建立一个随机投影矩阵对多尺度图像特征进行降维,并通过朴素贝叶斯分类器进行分类,其高效性引起了学者的广泛关注<sup>[9-13]</sup>。然而该算法依然存在以下问题:1)缺乏对目标观测信息是否有效的判断;2)由于遮挡等原因产生更新错误而导致目标漂移。

针对上述问题,本文充分利用文献[8]提取特征的高效性,从多尺度图像特征空间提取特征信息完成目标表征之后,由随机蕨分类器进行分类并通过一种特征置信度量策略来确定在线更新速率,最后由目标在线模型判断输出置信度最高的样本,此外还建立了一种遮挡反馈机制来决定是否更新目标在线模型,一旦判断目标出现遮挡,目标在线模型停止更新,分类器学习速率迅速降低,同时逐步扩大搜索半径以在后续帧中定位到目标。

## 2 实时压缩感知跟踪

### 2.1 随机投影矩阵提取特征

随机投影矩阵  $R$  的作用是将高维图像特征  $X$  压缩成低维特征  $V$  用以表征图像,并能以最小重构误差将图像还原<sup>[14]</sup>:

$$V = RX, \quad (1)$$

式中  $R \in \mathbb{R}^{n \times m}$ ,  $X \in \mathbb{R}^m$ ,  $V \in \mathbb{R}^n$ , 且  $n \ll m$ , 其中矩阵的选取至关重要,为了使图像以最小误差重构,文献[8]选取的随机投影矩阵为

$$r_{ij} = \sqrt{s} \times \begin{cases} 1, & p = 1/(2s) \\ 0, & p = 1 - 1/s \\ -1, & p = 1/(2s) \end{cases}, \quad (2)$$

式中  $r_{ij} \sim N(0, 1)$ ,  $s$  在 2 或 3 之间随机选择,因此随机投影矩阵中的大量元素都是零,这样就大大减少了计算量。

给定样本  $Z \in \mathbb{R}^{w \times h}$ , 它的高维特征  $X$  被定义为一系列的多尺度矩形滤波器,

$$h_{ij}(x, y) = \begin{cases} 1, & 1 \leq x \leq i, 1 \leq y \leq j \\ 0, & \text{otherwise} \end{cases}, \quad (3)$$

式中  $i$  和  $j$  是矩形滤波器的宽和高,  $w$  和  $h$  是样本窗口的宽和高,通过(1)式将高维特征压缩成低维特征  $V \in \mathbb{R}^n$ 。

### 2.2 朴素贝叶斯分类器

对于每个样本  $Z \in \mathbb{R}^{w \times h}$  而言,其低维特征表示为  $V = (v_1, \dots, v_n)^T$ , 假定样本标签为  $y \in \{1, 0\}$ , 1 代表正样本, 0 代表负样本。文献[8]假设低维特征中的每个元素是独立的,因此可以通过朴素贝叶斯分类器来建模:

$$H(V) = \log \left[ \frac{\prod_{i=1}^d p(v_i | y = 1) p(y = 1)}{\prod_{i=1}^d p(v_i | y = 0) p(y = 0)} \right] = \sum_{i=1}^n \log \left[ \frac{p(v_i | y = 1)}{p(v_i | y = 0)} \right]. \quad (4)$$

假设随机投影矩阵压缩后的低维特征服从高斯分布<sup>[8]</sup>, 则有:

$$p(v_i | y = 1) \sim N(\mu_i^1, \sigma_i^1), \quad p(v_i | y = 0) \sim N(\mu_i^0, \sigma_i^0). \quad (5)$$

对每一帧图像中的候选样本计算其  $H(V)$  值, 选择最大的候选样本作为最终跟踪结果, 再在其周围采样新的正样本, 计算其特征值并对(5)式进行更新:

$$\mu_i^1 \leftarrow \lambda \mu_i^1 + (1 - \lambda) \mu^1, \quad \sigma_i^1 \leftarrow \sqrt{\lambda (\sigma_i^1)^2 + (1 - \lambda) (\sigma^1)^2 + \lambda (1 - \lambda) (\mu_i^1 - \mu^1)^2}. \quad (6)$$

式中  $\mu^1 = \frac{1}{n} \sum_{k=0|y=1}^{n-1} v_i(k)$ ,  $\sigma^1 = \sqrt{\frac{1}{n} \sum_{k=0|y=1}^{n-1} (v_i(k) - \mu^1)^2}$ , 按照同样的方式对  $(\mu_i^0, \sigma_i^0)$  进行更新,  $\lambda$  为学习率, 越大代表更新速度越慢, 原始算法中固定为 0.85, 而一旦目标被遮挡则会出现更新错误, 当出现遮挡时间较长时分类器分类能力将不断减弱, 最终导致目标漂移。

### 2.3 目标定位与分类器更新

输入第  $t$  帧图像, 给定搜索半径  $\gamma$ , 以上一帧目标位置  $l_{t-1}$  为基准, 采样得到的样本位置集合为

$D^{\gamma} = \{z \mid \|l(z) - l_{t-1}\| < \gamma\}$ , 并对得到的样本通过(1)式提取出对应的低维特征  $V(z) = (v_1, \dots, v_n)^T$ 。

用(4)式计算样本集合  $D^{\gamma}$  中每一个样本低维特征  $V(z) = (v_1, \dots, v_n)^T$  的分类器响应值  $H(V)$ , 其中分类器响应值最大的即为当前帧的目标位置  $l_t$ 。

以当前帧的目标位置  $l_t$  为基准, 给定搜索半径  $\alpha$ ,  $\beta$  和  $\varsigma$ , 采样得到正样本集合  $D^{\alpha} = \{z \mid \|l(z) - l_t\| < \alpha\}$  和负样本集合  $D^{\varsigma, \beta} = \{s < z \mid \|l(z) - l_t\| < \beta\}$ , 再通过(1)式提取两个集合中样本对应的低维特征, 利用(6)式对分类器参数进行更新。

### 3 本文算法

从以上的分析可以看出, CT算法存在两个问题: 1) 假设所有低维特征是相互独立的, 忽略了特征之间的空间位置关系; 2) 学习率在整个跟踪过程中保持不变, 缺乏对遮挡等因素的反馈机制容易出现更新错误导致目标漂移。针对上述问题, 充分利用CT算法基于压缩感知理论提取特征的高效性, 并对其随机分组后交由随机蕨分类器进行分类, 由目标在线模型判断输出置信度最高的样本作为跟踪结果。算法框架图如图1所示。

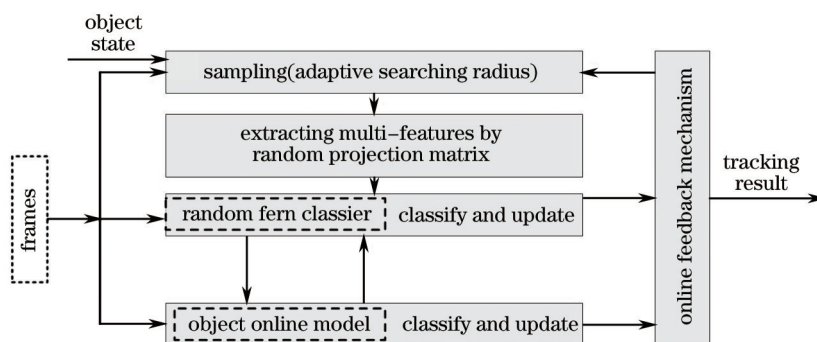


图1 算法框架图

Fig.1 Framework of the proposed diagram

#### 3.1 在线随机蕨分类器

随机蕨分类器是组合分类器的一种具体体现, 假设随机蕨分类器由不同的分类器构成, 每个分类器对未知样本做出独立判断, 则组合后的分类器的错误率将显著降低<sup>[15]</sup>。在CT算法中, 低维特征中的每一个元素都看成是独立的, 但样本中提取出的类 Haar 特征存在空间位置上的对应关系, 因此CT算法在减少计算量的同时也丢失了很多空间位置信息, 为此本文算法的解决方案是将特征集合  $V = (v_1, \dots, v_n)^T$  随机划分成大小为  $d$  的  $m$  个子集,

$$V_k = \{v_{\sigma(k,1)}, v_{\sigma(k,2)}, \dots, v_{\sigma(k,d)}\}, k = 1, \dots, m, \quad (7)$$

式中  $d = n/m$ ,  $\sigma(k, j)$  是在  $1, \dots, n$  范围内的随机数,  $V_k$  代表第  $k$  个蕨。每个蕨都对样本做出独立的判断, 可以通过朴素贝叶斯方法来建模:

$$H(V_k) = \log \left[ \frac{\prod_{i=1}^d p(v_{\sigma(k,i)} | y=1) p(y=1)}{\prod_{i=1}^d p(v_{\sigma(k,i)} | y=0) p(y=0)} \right] = \sum_{i=1}^d \log \left[ \frac{p(v_{\sigma(k,i)} | y=1)}{p(v_{\sigma(k,i)} | y=0)} \right], \quad (8)$$

$$y_k = \begin{cases} 1, & \text{if } H(V_k) > 0 \\ 0, & \text{otherwise} \end{cases} \quad (9)$$

给定一个判断阈值, 综合所有蕨的判断结果, 给输入样本打上标签  $Y \in \{1, 0\}$ , 1代表正样本, 0代表负样本, 表达式为

$$Y = \begin{cases} 1, & \text{if } \sum_{k=1}^m \alpha_k y_k > T_{fern} \\ 0, & \text{otherwise} \end{cases} \quad (10)$$

式中  $\alpha_k$  为每个蕨在综合判断中的权重, 设为  $1/m$ ,  $T_{fern}$  是判断阈值, 同时对每个蕨对应的特征参数  $\mu_i^l$  和  $\sigma_i^l$  都通过(6)式进行更新。

CT算法中学习率是固定不变的, 一旦出现更新错误将导致目标漂移, 基于高斯核密度估计提出一种特征置信度判断策略来自适应学习率的更新,

$$G(\mu_{t+1,i}) = \exp\left[-\frac{|\mu_{t+1,i} - \mu_{t,i}|^2}{(\sqrt{2}\sigma_{t,i})^2}\right], \lambda_{t+1,i} = \lambda_{t,i} G(\mu_{t+1,i}), \quad (11)$$

式中  $\mu_{t,i}$  和  $\sigma_{t,i}$  分别代表第  $t$  帧图像第  $i$  个特征分布的均值和方差,  $\lambda_{t,i}$  是对应于第  $t$  帧图像第  $i$  个特征的学习率, 初始化为  $\lambda_{1,i} = 0.85 (i = 1, \dots, n)$ 。当特征均值发生较大变化时, 学习率较小, 减少了更新错误的风险。

这样就构成了一个随机蕨分类器, 将通过随机蕨分类器的正样本交由下一步目标在线模型判断。其分类过程如图2所示。

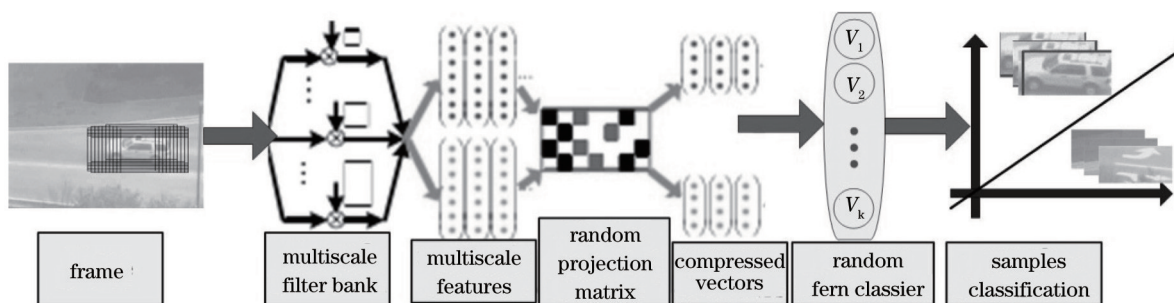


图2 随机蕨分类器工作过程

Fig.2 Working process of the random fern classiser

### 3.2 目标在线模型

目标在线模型  $M$  是包含有正负样本的数据结构,  $M = \{p_1^+, p_2^+, \dots, p_m^+, p_1^-, p_2^-, \dots, p_n^-\}$ , 其中  $p^+$  和  $p^-$  分别代表正负样本图像块。正样本按照时间顺序加入到  $M$  中,  $p_1^+$  由第一帧初始化产生,  $p_m^+$  则代表最新加入到  $M$  中的正样本图像块。为方便处理, 所有正负样本图像块都由采样样本归一化到  $15 \text{ pixel} \times 15 \text{ pixel}$ , 这样两个样本图像块的相关相似度定义为

$$S(p_i, p_j) = 0.5[f_{\text{NCC}}(p_i, p_j) + 1], \quad (12)$$

式中  $f_{\text{NCC}}(p_i, p_j)$  为两个图像块之间的归一化互相关系数<sup>[7]</sup>。

目标在线模型的构建与更新: 输入一个图像块  $p$ , 定义其正最近邻相似度为  $S^+(p, M) = \max_{p_i^+ \in M} S(p, p_i^+)$ , 负最近邻相似度为  $S^-(p, M) = \max_{p_i^- \in M} S(p, p_i^-)$ , 则该样本的置信度定义为

$$f_{\text{conf}} = \frac{1 - S^-(p, M)}{1 - S^+(p, M) + 1 - S^-(p, M)}. \quad (13)$$

当图像块的标签为1且  $f_{\text{conf}} < T_{\text{nnP}}$  (实验中取0.7)时, 说明目标外观发生了一定的变化, 所以将其作为可靠的正样本图像块加入到在线模型中, 当图像块的标签为0且  $f_{\text{conf}} > T_{\text{nnP}}$  (实验中取0.5)时, 说明该负样本图像块与正样本具备一定的相似性, 干扰性较强, 能够提高目标在线模型的区分能力, 所以将其作为可靠的正样本图像块加入到在线模型中。图3给出了 panda.mpg 视频处理完成后所生成的在线模型中归一化正负样本图像块。

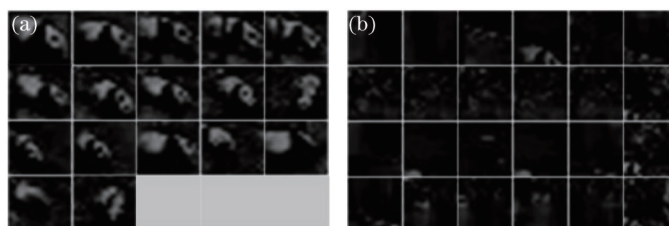


图3 目标在线模型。(a)正样本本图像块; (b)负样本本图像块

Fig.3 Online object model. (a) positive image patches; (b) negative image patches

目标定位:首先将通过随机蕨分类器的  $k$  个样本归一化为  $15 \text{ pixel} \times 15 \text{ pixel}$  的图像块,由(13)式计算每一个样本的  $f_{\text{conf}}$ ,则最终的目标位置  $l(z^*)$  定义为

$$l(z^*) = \underset{z}{\text{argmax}}_{f_{\text{conf}}(z) > T_{\text{nn}}} f_{\text{conf}}(i), i = 1, 2, \dots, k. \quad (14)$$

如果通过随机蕨分类器的  $k$  个样本都不满足  $f_{\text{conf}}(i) > T_{\text{nn}}$ ,则抛出目标丢失的警告,下一部分来讨论对于目标遮挡或完全丢失的反馈机制及处理方案。

### 3.3 目标遮挡反馈机制

实际应用中,大量因素影响着跟踪算法的性能,而其中目标被遮挡的情况最常见。为应对这些干扰因素,引入了遮挡反馈机制来提升算法的整体性能。

在CT算法中利用每一帧的跟踪结果对分类器进行更新,且学习率在整个处理过程中固定不变,假定每一帧跟踪结果是准确的,一旦目标长时间被遮挡,整个分类器将会出现跟踪错误而逐渐失去判别能力,导致目标重新出现时无法输出正确的跟踪结果。本文引入一种遮挡反馈机制:通过评估当前的跟踪状态来决定目标是否被遮挡,以及分类器是否需要更新、目标在线模型是否需要更新、搜索窗口半径是否需要扩大。

对于第  $t$  帧图像,定义  $S_t$  代表最终跟踪结果的目标状态,  $S_t = 0$  代表目标被遮挡,分类器将不再更新,目标在线模型保持不变,搜索窗口半径  $\gamma_t$  应当增大,增大比例系数为  $\tau$ ;  $S_t = 1$  代表跟踪结果可靠,目标状态稳定,随机蕨分类器学习率  $\lambda_t$  将按照(11)式继续更新,目标在线模型根据(13)式进行更新,搜索半径保持初始值  $\gamma_1$  不变。定义  $C_t$  代表最终跟踪结果的随机蕨分类器置信度,由(8)式得到  $C_t = \sum_{k=1}^m H(V_k)$ ,定义前后两帧跟踪结果置信度变化量为  $\Delta C_t = C_t - C_{t-1}$ ,定义  $t_{\text{tol}}$  为置信度变化量的容忍度(实验中取 10),定义  $f_{\text{conf}}$  代表最终跟踪结果通过目标在线模型的置信度。则对目标是否出现遮挡的情况如表 1 所示。

表 1 目标遮挡反馈机制

Table 1 Feedback mechanism for object occlusion

Index	$S_t$	$\lambda_t$	$\gamma_t$	Determining condition
1	0	0	$\gamma_{t-1}\tau$	$C_t < 0$
2	0	0	$\gamma_{t-1}\tau$	$\Delta C_t = C_t - C_{t-1} < t_{\text{tol}}$
3	0	0	$\gamma_{t-1}\tau$	$f_{\text{conf}} < 0.5$
4	1	$\lambda_t G(\mu_t)$	$\gamma_t$	$C_t > 0 \& \Delta C_t > t_{\text{tol}} \& f_{\text{conf}} \geq 0.5 \& S_{t-1} = 0$
5	1	$\lambda_{t-1} G(\mu_t)$	$\gamma_t$	$C_t > 0 \& \Delta C_t > t_{\text{tol}} \& f_{\text{conf}} \geq 0.5 \& S_{t-1} = 1$

### 3.4 算法描述

算法分为第一帧初始化和后续帧跟踪两部分。

输入第 1 帧图像:

1) 参数初始化:低维特征数目、搜索窗口半径、随机蕨分类器的参数、目标在线模型的置信度阈值等。生成随机矩阵中非零元素对应的特征位置并将低维特征随机平均分配给随机蕨分类器的每个蕨(在后续帧处理中保持不变)。

2) 分类器和目标在线模型构建:按照 2.3 节中 CT 算法的方式采样生成正负样本集合  $D^{\alpha}$  和  $D^{\beta}$ ,再训练随机蕨分类器和目标在线模型。

输入第  $t$  帧图像:

1) 目标定位:输入目标遮挡反馈机制产生的搜索窗口半径和上一帧目标位置,采样生成候选样本集合  $D^y$ ,将通过随机蕨分类器和目标在线模型的样本位置作为跟踪结果,否则抛出目标丢失的警告。

2) 随机蕨分类器更新:按照(6)式对每个蕨对应的特征参数  $\mu_i^j$  和  $\sigma_i^j$  进行更新,按照(11)式对每个特征对应的学习率  $\lambda_{i,j}$  进行更新。

3) 目标在线模型更新:将通过随机蕨分类器的样本按照(13)式计算其置信度,将满足 3.2 中阈值条件的样本加入到目标在线模型。

4) 目标遮挡反馈机制:依据表 1 评估最终的跟踪结果来决定目标是否被遮挡,以及分类器是否更新、目标在线模型是否更新、搜索窗口半径是否扩大。

5) 读取下一帧图像,返回步骤 1)。

## 4 实验与结果分析

为验证本文算法的有效性,选用不同场景下具有代表性的三组测试视频 Running Car<sup>[6]</sup>、Suv<sup>[7]</sup>、David indoor<sup>[8]</sup>, 与原始 CT<sup>[8]</sup>和 TLD<sup>[7]</sup>跟踪算法进行对比实验。实验平台采用 Intel Core(TM)2 Duo CPU 主频 2.5GHz 的 PC 机,利用 OpenCV2.4.9 在 VS2010 上进行调试。

### 4.1 参数设置

压缩后的低维特征数目  $n = 130$ , 随机蕨分类器中蕨的数目  $m = 10$ , 每个蕨包含的特征数目  $d = 13$ , 分类器学习率初始值  $\lambda_1 = 0.85$ , 搜索窗口半径初始值  $\gamma_1 = 25$ , 窗口增大比例系数  $\tau = 1.5$ , 正样本采样半径  $\alpha = 4$ , 负样本内外采样半径  $s = 8$  和  $\beta = 30$ , 随机蕨分类器判断阈值  $T_{ferm} = 0.5$ 。CT<sup>[8]</sup>和 TLD<sup>[7]</sup>跟踪算法采用原文中的参数设置。

### 4.2 结果分析

实验采用两个评估标准对算法的性能进行测试:1) 跟踪成功率(SR), 定义  $f_{score} = \frac{R_T \cap R_C}{R_T \cup R_C}$  代表每一帧跟踪结果的得分, 其中  $R_T$  代表跟踪结果区域,  $R_C$  代表目标真实区域(人工标注), 如果  $f_{score} > 0.5$  则认为跟踪结果是正确的;2) 中心位置误差, 为  $\sqrt{(x_T - x_C)^2 + (y_T - y_C)^2}$ , 其中  $(x_T, y_T)$  是跟踪结果的目标位置,  $(x_C, y_C)$  是目标真实位置, 定义误差低于 20 pixel 为正确的跟踪结果。表 2 给出了三种算法在测试视频中的成功率对比, 图 4 给出了不同算法的中心位置误差对比。

表 2 不同算法跟踪成功率对比  
Table 2 Comparison of success rate for different tracking methods

Sequence	Proposed /%	CT /%	TLD /%
Running Car	100	98	63
Suv	98	33	92
David indoor	94	89	96

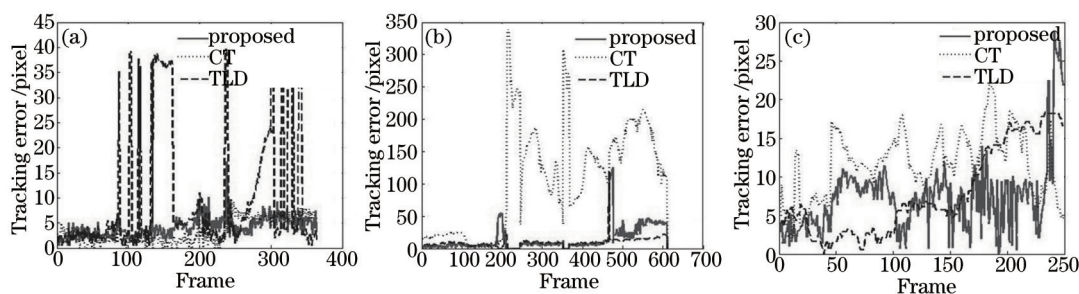


图 4 视频序列跟踪误差。(a) Running Car 序列; (b) Suv 序列; (c) David indoor 序列

Fig.4 Tracking error of video sequences. (a) Tracking error of Running Car; (b) tracking error of Suv; (c) tracking error of David indoor

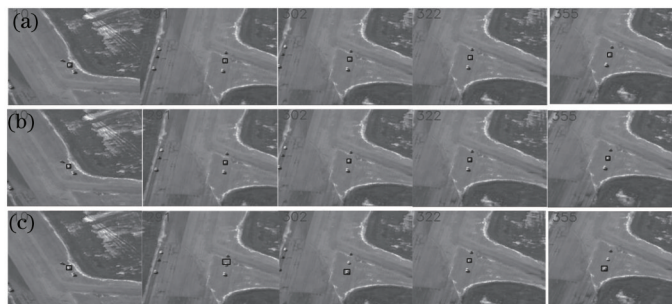


图 5 Running Car 视频序列跟踪结果。(a) 本文算法; (b) CT 法; (c) TLD 法

Fig.5 Tracking results of Running Car video sequence. (a) Proposed method; (b) CT method; (c) TLD method

图 5 显示了本文算法的抗相似目标干扰能力。任务是跟踪大小为 22 pixel×18 pixel 的中间那辆白色小

车, Running Car 视频序列中由于大场景下目标尺寸较小, 纹理特征信息较弱, 所以前后三辆小车极为相似。从 Running Car 视频序列跟踪结果来看, TLD 算法中的跟踪模块采用的中值流跟踪, 在该视频目标与背景纹理特征变化不大时容易出现跟踪漂移, 所以在第 291 帧时跟踪框偏离了目标, 而检测模块采取的全局扫描窗口策略则很容易使跟踪器重新定位到相似目标上, 进而导致跟踪错误, 出现了第 302、355 帧等跟踪到第三辆车的错误结果, 图 4(a) 跟踪误差曲线中 TLD 算法多处出现跟踪误差剧烈增大, 也说明了跟踪了错误的目标。而本文算法由于采取局部采样策略及低维特征的多样性, 和 CT 算法在整个跟踪过程中表现稳定, 误差均保持在 20 pixel 以内, 跟踪成功率达到 100%。

图 6 显示了本文算法的抗遮挡能力。目标大小为 108 pixel×53 pixel。Suv 视频序列中多次出现遮挡: 第 185~245 帧长时遮挡、第 345~365 帧和第 450~465 帧短时遮挡。CT 算法无法做出遮挡判断, 图 4(b) CT 算法的误差曲线反应了 200 帧以后分类器完全丧失判断能力, 当目标第 246 帧重新出现时无法继续完成跟踪任务。而本文算法采用的随机蕨分类器利用了类 Harr 特征的空间信息, 每个蕨分类器对应一定空间分布的特征, 分类过程中单独做出判断, 最后综合各个蕨分类器的判断结果交由遮挡反馈机制, 一旦判定出现完全遮挡则抛出警告(如图 6 第 192、451 帧左上侧的红点), 停止对分类器和目标在线模型的更新并逐步扩大搜索半径, 因此在目标重新出现时能将其准确定位并继续跟踪(如图 6 中的第 246、366、465 帧), 跟踪误差基本保持在 20 pixel 以内, 跟踪成功率达到 98%。

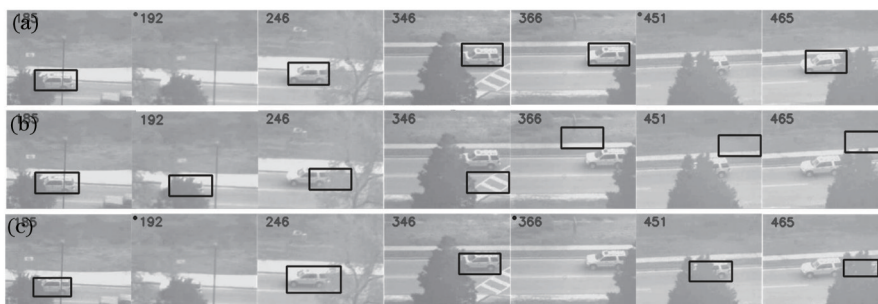


图 6 Suv 视频序列跟踪结果。(a) 本文算法; (b) CT 法; (c) TLD 法

Fig.6 Tracking results of Suv video sequence. (a) Proposed method; (b) CT method; (c) TLD method

图 7 显示了本文算法的抗光照变化和物体形变能力。目标大小为 73 pixel×70 pixel, David indoor 视频序列中光照由弱到强, 人体在第 170 帧发生转动时目标形变较大。由于本文算法采用目标在线模型, 能够不断学习到目标形变并将其加入到目标在线模型中, 因此在整个跟踪过程中能稳定跟踪目标, 平均跟踪误差比 CT 算法低 4.43 pixel, 与 TLD 算法相当。但是在后期背景干扰较大的情况下, 目标在线模型出现些许更新错误, 从图 4(c) 中发现共 15 帧的跟踪误差超过 20 pixel, 跟踪成功率为 94%。



图 7 David indoor 视频序列跟踪结果。(a) 本文算法; (b) CT 法; (c) TLD 法

Fig.7 Tracking results of David indoor video sequence. (a) Proposed method; (b) CT method; (c) TLD method

对于目标跟踪实际应用, 算法的实时性也是重要的评估指标, 表 3 给出了三种算法的处理速度对比, 本文算法采用随机投影矩阵提取低维目标特征并分组交由随机蕨分类器分类, 在保证特征的多样性的同时减少了计算量, 但建立目标模型并在线更新使得处理速度较 CT 算法有所降低, 依然保持在 30~50 frame/s 左右, 因此完全满足实时性的要求。

表 3 三种算法处理速度对比(帧/秒)

Table 3 Comparison of three algorithms processing speed (frame /s)

Sequence	Proposed method	CT method	TLD method
Running Car	28.71	51.69	12.33
Suv	41.01	73.35	20.78
David indoor	50.51	52.10	17.21

## 5 结 论

针对现有在线学习跟踪算法存在的两个问题,提出了一种稳定且实时的在线学习跟踪算法。基于压缩感知理论提取低维目标特征并分组交由随机森林分类器分类,大大减少了算法计算量,在保证特征的多样性的提升了算法的实时性,构建目标在线模型实时更新建立了一种遮挡反馈机制,使得目标被遮挡后及光照变化时依然能够准确定位,提升了算法的稳定性。通过三组实验表明,本文方法在有相似目标干扰、目标被长时或短时遮挡、光照变化以及目标形变的情况下都能得到较好的跟踪结果,处理速度保持在 30~50 frame/s 左右,完全满足实时性的要求。此外本文算法可以与现有的尺度变化处理机制相结合,来自适应目标的尺度变化,这也是后续的研究重点。

## 参 考 文 献

- 1 Wu Y, Lim J, Yang M H. Online object tracking: A benchmark[C]. Computer Vision and Pattern Recognition, Conference on. IEEE, 2013, 2411-2418.
- 2 Ross D A, Lim J, Lin R S, *et al.*. Incremental learning for robust visual tracking[C]. International Journal of Computer Vision, 2008, 77 (1-3): 125-141.
- 3 Mei X, Ling H. Robust visual tracking and vehicle classification via sparse representation[C]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2011, 33(11): 2259-2272.
- 4 Li H, Shen C, Shi Q. Real-time visual tracking using compressive sensing[C]. Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on. IEEE, 2011: 1305-1312.
- 5 Grabner H, Leistner C, Bischof H. Semi-Supervised On-Line Boosting for Robust Tracking[M]. Heidelberg: Springer, 2008: 234-247.
- 6 Babenko B, Yang M H, Belongie S. Robust object tracking with online multiple instance learning[J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2011, 33(8): 1619-1632.
- 7 Kalal Z, Mikolajczyk K, Matas J. Tracking-learning-detection[J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2012, 34(7): 1409-1422.
- 8 Zhang K, Zhang L, Yang M H. Real-Time Compressive Tracking[M]. Heidelberg: Springer, 2012: 864-877.
- 9 Wu Y, Jia N, Sun J. Real-time multi-scale tracking based on compressive sensing[J]. The Visual Computer, 2014, 31(4): 471-484.
- 10 Zhu Q, Yan J, Deng D. Compressive tracking via oversaturated sub-region classifiers[J]. Computer Vision, IET, 2013, 7(6): 448-455.
- 11 Jing J, Xu G, Lei B, *et al.*. Using appearance re-matching to improve real-time compressive tracking[C]. Proceedings of the Fifth International Conference on Internet Multimedia Computing and Service. ACM, 2013: 28-33.
- 12 Zhu Qiuping, Yan Jia, Zhang Hu, *et al.*. Real-time tracking using multiple features based on compressive sensing[J]. Optics and Precision Engineering, 2013, 21(2): 437-444.  
朱秋平, 颜 佳, 张 虎, 等. 基于压缩感知的多特征实时跟踪[J]. 光学 精密工程, 2013, 21(2): 437-444.
- 13 Mao Zheng, Yuan Jian jian, Wu Zhen rong, *et al.*. Real-time compressive tracking based on online feature selection[J]. Optics and Precision Engineering, 2014, 22(3): 730-736.  
毛 征, 袁建建, 吴珍荣, 等. 基于在线特征选择的实时压缩跟踪[J]. 光学 精密工程, 2014, 22(3): 730-736.
- 14 Yank Sa, Yang Chunling. Image registration algorithm based on sparse random projection and scale-invariant feature transform[J]. Acta Optica Sinica, 2014, 34(11): 1110001.  
杨 飒, 杨春玲. 基于压缩感知与尺度不变特征变换的图像配准算法[J]. 光学学报, 2014, 34(11): 1110001.
- 15 Özuysal M, Calonder M, Lepetit V, *et al.*. Fast keypoint recognition using random ferns[J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2010, 32(3): 448-461.
- 16 Guo Pengyu, Su Ang, Zhang Hongliang, *et al.*. Online mixture of random naive Bayes tracker combined texture with shape feature[J]. Acta Optica Sinica, 2015, 35(3): 0315002.  
郭鹏宇, 苏 昂, 张红良, 等. 结合纹理和形状特征的在线混合随机朴素贝叶斯视觉跟踪器[J]. 光学学报, 2015, 35(3): 0315002.

栏目编辑: 苏 岑