

基于核相关滤波的多目标追踪

刘欢, 李春庚*, 安居白, 魏帼, 任俊丽

大连海事大学信息科学技术学院, 辽宁 大连 116026

摘要 为了克服核相关滤波(KCF)只根据目标外观模型追踪时准确性低的不足,融入运动模型,计算了检测目标框和预测目标框的交并比(IOUS)。通过匈牙利算法,确定了目标间的最优关联。KCF 和 IOU 模型都具有快速响应的特点,因此算法可满足在线处理数据的要求。在公开的 2DMOT2015、MOT16 数据集上进行实验,将所提方法与其他优秀方法相比,在保证 30 frame/s 以上处理速度的同时,追踪准确性提高 10% 以上。

关键词 机器视觉; 多目标跟踪; 核相关滤波; 交并比

中图分类号 TP391.41

文献标识码 A

doi: 10.3788/LOP56.121501

Multiple Object Tracking Based on Kernelized Correlation Filter

Liu Huan, Li Chungeng*, An Jubai, Wei Guo, Ren Junli

Information Science and Technology College, Dalian Maritime University, Dalian, Liaoning 116026, China

Abstract A motion model is used to overcome the shortcoming of the low tracking accuracy of a kernelized correlation filter (KCF) based only on a target appearance model. The intersection-over-union (IOU) between the detection target bounding box and the predicted target bounding box was calculated. The optimal correlation among the targets was determined using the Hungarian algorithm. Both the KCF and IOU models are characterized by fast responses; therefore, the algorithm has the ability to process data online. The experiments were conducted on the public 2DMOT2015 and MOT16 datasets. Compared with the other state-of-the-art method, the tracking accuracy of the proposed method is higher than 10% while ensuring a processing speed of 30 frame/s or faster.

Key words machine vision; multiple object tracking; kernelized correlation filter; intersection-over-union

OCIS codes 150.0155; 100.4996; 100.4999; 100.2000

1 引言

随着计算机视觉技术的发展,人们对于视频内容理解分析的要求越来越高,其中目标追踪分支在机器人自动避障导航、无人车自动驾驶、行人行为监控等方向都有着很广泛的应用前景。针对不同的应用方向,要解决的关键问题有所不同,对于追踪的准确性或者追踪的实时性也有着不同的要求。

对于不同的应用场景,业界一直在寻找一种性能比较均衡且具有一定通用性的方法,去解决目标遮挡、相似目标干扰、天气变化、相机运动等问题。为了比较不同算法的性能,国外有学者提供了 2DMOT2015、MOT16 数据集^[1]供大家研究和使

用。数据集由不同序列组成,不同序列体现了不同的应用场景,希望通过此数据集可以统一评判不同算法的追踪效果。

由于追踪的速度和准确性通常是相互矛盾的,因此国内外学者提出的多目标跟踪(MOT)算法,可以根据侧重点分为准确性提升和追踪速度提升两种方法。提升准确性的算法有多假设追踪模型^[2](MHT)、马尔可夫决策模型^[3](MDP)、卷积网络的在线多目标追踪^[4](AP_RCNN)、连续最小能量法^[5](CEM)、多特征融合的相关滤波追踪(MAPRF)^[6]等;提升追踪速度的算法有交并比^[7](IOU-tracker)、简单在线实时追踪^[8](SORT)、深度关联简单在线实时追踪^[9]等。从这些算法可以看到,对

收稿日期: 2018-12-04; 修回日期: 2019-01-04; 录用日期: 2019-01-14

基金项目: 国家自然科学基金(61471079)

* E-mail: li_chungeng@dlnu.edu.cn

于多目标追踪问题,仅利用目标位置信息建立运动模型追踪速度快,但追踪的准确性较差;当使用外观模型进行追踪处理时,追踪准确性高,但会使处理速度减慢。

由于相关滤波算法在单目标追踪领域的成功应用,算法把实数域的数据处理过程转化为频率域处理,处理速度得到极大地提升。文献[10]中提出核相关滤波算法,该算法在相关滤波基础上,采用方向梯度直方图(HOG)^[11]方法进行特征提取,并构建能够处理多通道的核函数,实现目标的稳定可靠追踪。文献[12]在核相关滤波(KCF)基础上添加尺度更新机制,解决了目标尺度变化追踪漂移等难题,但处理速度变慢。本文利用 KCF 处理快速的优势,将 KCF 算法应用于多目标追踪过程的外观模型,融合 IOU 和匈牙利算法实现数据关联,提出一种目标相似比较方法以解决目标重现问题,具体工作如下:

1) 由于原数据集提供的行人位置信息不够准确,存在误检、漏检等情况,因此利用 Faster-RCNN (VGG16)^[13] 目标检测网络对图像序列重新检测。

2) 利用 KCF 建立目标的外观模型。由于 KCF 使用了 HOG^[11] 特征、循环矩阵傅里叶对角化、多通道高斯核函数等技巧,目标追踪具有很好的准确性和处理数据的快速性,因此,将其融合到多目标追踪中,针对每个目标建立一个追踪器,从而进行目标训练和检测。

3) 用目标相似性判断方法解决目标重现问题。当 2 个目标图像平均像素误差小于特定阈值时,认为是同一个目标;同时设定一个最大丢失时间,当目标丢失时间超过此阈值则认为目标彻底消失,追踪结束。

2 核相关滤波追踪

KCF 算法^[10]对追踪样本通过循环位移方式,得到稠密样本构成循环矩阵,利用循环矩阵傅里叶对角化的性质进行矩阵分解,简化矩阵计算,使得目标追踪的训练和检测过程大幅加速;为了提高追踪效果,KCF 算法使用更具信息表达能力的多通道 HOG 特征代替单通道的灰度特征,并引入能够处理多通道的高斯核函数来处理 HOG 等特征。

与传统的模板匹配方法相比,追踪效果和追踪速度都得到很大的提高。KCF 算法通过使用核函数方式来提高算法判别能力,常用的核函数有线性、多项式、高斯核等,其中高斯核函数表达式^[10]为

$$k^{uv} = \exp \times$$

$$\left\{ -\frac{1}{\sigma^2} \left[\| \mathbf{u} \|^2 + \| \mathbf{v} \|^2 - 2\mathcal{F}^{-1} \left(\sum_c \hat{\mathbf{u}}_c^* \cdot \hat{\mathbf{v}}_c \right) \right] \right\}, \quad (1)$$

式中: σ 为高斯核设定的标准差,需要根据实验设定; \mathcal{F}^{-1} 表示逆傅里叶变换; \mathbf{u} 和 \mathbf{v} 为相同大小的任意样本; $\hat{\mathbf{u}}_c$ 和 $\hat{\mathbf{v}}_c$ 为对应样本在第 c 个通道的傅里叶变换形式; $\hat{\mathbf{u}}_c^*$ 为 $\hat{\mathbf{u}}_c$ 的复数共轭矩阵;“ \cdot ”为点乘操作。

KCF 追踪器的训练过程为

$$\hat{\boldsymbol{\alpha}} = \frac{\hat{\mathbf{y}}}{\hat{\mathbf{k}}^{xx} + \lambda}, \quad (2)$$

式中: $\boldsymbol{\alpha}$ 表示追踪器模板参数; \mathbf{x} 为训练目标的 HOG 特征; \mathbf{y} 的大小与 \mathbf{x} 相同,元素值符合二维高斯分布; λ 为正则化参数; \mathbf{k}^{xx} 为训练目标 \mathbf{x} 与自己的核相关结果。当样本大小为 $(n \times n)$ 时,(2)式时间复杂度为 $O(n^2 \log n)$ 。

KCF 追踪器的检测过程为

$$\hat{f}(\mathbf{z}) = \hat{\mathbf{k}}^{xz} \cdot \hat{\boldsymbol{\alpha}}, \quad (3)$$

式中:形如 $\hat{\mathbf{h}}$ 表示该参数的离散傅里叶变换向量; $\hat{f}(\mathbf{z})$ 表示待检测区域的检测响应结果; $\boldsymbol{\alpha}$ 为(2)式的追踪器模板参数; \mathbf{x} 为训练目标的 HOG 特征; \mathbf{z} 为待检测区域的 HOG 特征; \mathbf{k}^{xz} 为训练目标和待检测目标的核相关结果。当样本大小为 $(n \times n)$ 时,(3)式计算时间复杂度为 $O(n^2 \log n)$ 。

将 KCF 算法应用到多目标追踪中,仅能实现外观模型关联。由于 KCF 不支持多尺度情况,因此检测到的目标框存在过大或者过小等不合理问题,在后续的目标追踪过程中易出现偏差;对于场景中目标密集情况,一个 KCF 追踪器的预测结果会和多个目标有交集,此时 KCF 不能做出合理的关联判断;在多目标追踪过程中,还会出现目标短暂消失、被建筑物遮挡的情况,KCF 不具备处理目标重现的能力。

3 改进的核相关滤波追踪方法

本文提出了使用 IOU 和匈牙利算法的方式解决 KCF 不支持多尺度带来追踪偏差以及目标关联问题。针对目标的非惯性运动会目标丢失若干帧后重现的情况,设计了目标相似性判断方法;然后分析了单个目标在视野中会出现的各种状态,以及状态间的转换关系;之后提出基于 KCF 实现的多目标追踪方法。

由于数据集提供的目标框标注不精准,存在较严重的误检和漏检等情况。为了解决相关问题,文献[13]中提出 Faster-RCNN 目标检测网络,文献[14]中针对暗通道有雾情况对 Faster-RCNN 进行改进,更好地解决了雾霾天气下的目标检测问题。因此,本文采用 Faster-RCNN 基本网络进行目标检测,该网络可以检测包括行人在内的 20 种类,研究重点仅关注行人目标。检测结果为多组数据,每组数据由 5 个数据构成 (x, y, w, h, c_{score}) , 分别表示为目标在图像中左上角的纵横坐标位置、目标的宽和高、检测结果的置信度, (x, y, w, h) 称为目标框 (bbox)。

3.1 目标关联策略

对于多目标追踪问题,前后帧目标数量大都是变化、不确定的,在进行数据关联时,要求采用一种全局最优的方式,将数据尽可能关联。为了解决该问题,采用 IOU 和匈牙利算法结合的方式。例如,

在 $k-1$ 帧时训练 p 个 KCF 追踪器,在 k 帧时出现 q 个目标,此时需要将这些目标一一计算 IOU 值,构成一个大小为 $q \times p$ 的 IOU 矩阵,利用匈牙利算法对这个矩阵进行优化,寻找出最佳关联结果。IOU 是用来计算 1 幅图像中 2 个目标框重叠部分所占比例的一个公式,即

$$S_{IOU}(a, b) = \frac{A_{area}(a) \cap A_{area}(b)}{A_{area}(a) \cup A_{area}(b)}, \quad (4)$$

式中: A_{area} 表示目标框面积; $A_{area}(a) \cap A_{area}(b)$ 为 a, b 两个目标框面积的交集; $A_{area}(a) \cup A_{area}(b)$ 为 a, b 两个目标框面积的并集。图 1 为数据集 ADL-Rundle-6 序列的第 50 帧画面,黄色矩形表示 KCF 的目标检测结果,定义为 $a = (x_t, y_t, w_t, h_t)$, 下标 t 表示追踪的意思;红色矩形表示 Faster-RCNN 的目标检测结果,定义为 $b = (x_d, y_d, w_d, h_d)$, 下标 d 表示检测意思。使用(4)式对 a, b 两个目标框进行 IOU 计算,判定追踪结果和检测结果的重叠率,计算式为

$$S_{IOU}(a, b) = \frac{(x_t + w_t - x_d) \times (y_t + h_t - y_d)}{w_t \times h_t + w_d \times h_d - (x_t + w_t - x_d) \times (y_t + h_t - y_d)}, \quad (5)$$

式中: S_{IOU} 的取值范围为 $0 \sim 1$, 当两个 bbox 完全重合时,说明两个目标完全匹配,此时结果为 1; 当两个 bbox 完全不相关时,取值为 0, 算法不会对其进行关联。通过计算 IOU 的方式,可以有效减少目标匹配计算过程,提高算法运行效率。

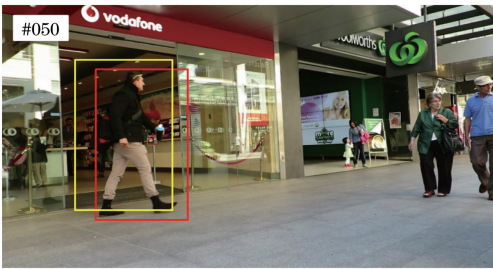


图 1 2DMOT2015 数据集 ADL-Rundle-6 序列的 IOU 展现
Fig. 1 IOU presentation of ADL-Rundle-6 sequence in 2DMOT2015 dataset

3.2 目标相似性判断方法

目标在追踪过程中会出现遮挡、消失后重现等情况,如果不进行目标外观判断就会把重现的目标当成新目标,从而导致目标追踪不准确。在保证算法具有快速处理能力的情况下,为了解决该问题,根据

$$d = \frac{\sum_{i'}^m \sum_{j'}^n \text{abs}(I_{i'j'}, P_{i'j'})}{m \times n} \quad (6)$$

进行计算,式中: I 和 P 分别为新出现的目标和丢失的目标灰度图像; i' 和 j' 表示图像横坐标索引和纵坐标索引;把 P 图像处理为 I 图像相同大小,记为 $m \text{ pixel} \times n \text{ pixel}$; $\text{abs}(x, y)$ 表示计算两个变量差值的绝对值函数。计算两个目标对应位置像素差的平均值,即为相似度距离。

图 2 来源于 2DMOT2015 数据集 ADL-Rundle-6 序列。图 2(a)、(b)为同一个目标在不同时间的截图,将目标处理为灰度图后进行像素差值计算,并得到图 2(c),此时图像误差平均值为 17.5855。图 2(c)中颜色越黑说明像素点差值越小,可以通过设置某个阈值来判定是否为一个目标。对



图 2 目标相似性判断。(a)第 4 帧;(b)第 21 帧;
(c)模板间变化
Fig. 2 Object similarity judgment. (a) 4th frame; (b) 21st frame; (c) template-to-template change

于此阈值的设定需要实验确定,如果设置较小则不能将目标关联,这是因为简单灰度图像包含部分背景信息,同时目标姿态变化也会带来影响;如果阈值设置过大,则会把不同的目标判定为同一个目标,导致追踪结果漂移严重。

3.3 目标转换状态过程

文献[3]中将目标在相机视野中归结为4种状态(active 状态、inactive 状态、tracked 状态、lost 状态)和状态间的7条变化路径,如图3所示。其中,active 为开始状态,根据 bbox 判断目标是否为行人,如果是,则根据路径 a_1 转移到 tracked 状态;如果不是,则直接由路径 a_2 标定为 inactive 状态,并终止该目标的追踪。同一个目标在视野中会出现漏检或者被遮挡的情况,以及被遮挡后目标重现等情况,如图3中的 a_3 、 a_4 、 a_5 、 a_6 、 a_7 路径。

在本文算法中,对于 Faster-RCNN (VGG16) 检测到的目标,选择行人部分,然后根据检测框的 c_{score} 值,执行 active 状态的策略。大于一定阈值的,判定为行人转为 tracked 状态;小于一定阈值的,判定为不可靠检测转为 inactive 状态,结束追踪。对于 lost 状态,利用目标相似性判断方法解决目标重现问题。

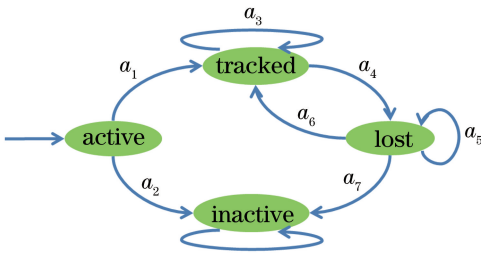


图3 目标流转状态和路径

Fig. 3 Status and path of target flow

3.4 算法流程

本文算法利用 KCF 算法解决多目标的外观模型问题,使用 IOU 关联数据,构建本文算法伪代码形式,如图4所示。算法从输入解释、输出解释和运行代码等3个方面进行说明。由于多目标追踪问题参数众多,在图4伪代码中,假定有 F 帧图像输入,每一帧有 N 个检测目标,变量 I 、 D 、 T 分别表示输入图像集合、检测框集合、追踪的结果;所以在第 i 帧的时候会有 $D_i = \{d_{i1}, d_{i2}, \dots, d_{iN}\}$ 个目标检测框和 $T_i = \{t_{i1}, t_{i2}, \dots, t_{iN}\}$ 个追踪结果; T_i 表示第 i 帧的追踪结果; $T_j.id$ 表示所有追踪器在第 j 帧的序号; T_{ij} 表示第 i 帧的第 j 个 KCF 追踪器; $T_{ji}.id$ 、 $T_{ji}.pos$ 、 $T_{ji}.m$ 分别为第 j 帧中第 i 个目标的序号、

Input	$I=\{I_1, \dots, I_F\}$ $D=\{D_1, \dots, D_F\}$
Output	$T=\{T_1, \dots, T_F\}$
Process	<pre> for $d_i \in D_i$: $T_{ji}.id = g_id++$ $T_{ji}.m = KCF_train(I_j, d_i)$ </pre>
	<pre> for $j=2$ to F: for $t_i \in T_{j-1}$: $T_{ji}.pos = KCF_detect(I_j, t_i)$ $iou = IOU(T_{ji}, D_j)$ $T_{ji}.id = Hungarian(iou)$ for $d_j \in D_j$: if $(T_{ji}.id == -1)$: $T_{pre} = find_target(T_j, j, K_{lost})$: $r_score, r_id = associate(T_{pre}^{obj}, T_{pre})$: if $r_score < K_{same}$: $T_{ji}.id = r_id$ else: $T_{ji}.id = g_id++$ $T_{ji}.m = KCF_train(I_j, d_i)$ </pre>

图4 KCF-IOU 追踪算法

Fig. 4 KCF-IOU tracking algorithm

位置和模板; K_{lost} 、 K_{same} 、 K_{score} 分别表示最大丢失时间、目标相似性阈值、行人置信度。

对 KCF-IOU 追踪算法的详细描述过程如下:

- 1) 假定有 F 帧图像输入,每一帧有 N 个检测目标。
- 2) 在第1帧中先根据 Faster-RCNN 检测到的 bbox 利用(2)式进行 KCF 训练,得到每个目标的 KCF 追踪器(tracker)。
- 3) 在后续每帧中先利用(3)式进行 KCF 检测,得到对应目标的位置,利用 KCF 检测到的 bbox 与 Faster-RCNN 检测到的 bbox 作 IOU 计算,对于多个目标相互关联,会得到一个 IOU 矩阵。
- 4) 利用匈牙利算法进行 IOU 矩阵分配,得到最优解。
- 5) KCF 检测到的 bbox 和 Faster-RCNN 检测到的 bbox 成功关联的标记为 tracked 状态。
- 6) KCF 检测到的 bbox 未能和 Faster-RCNN 检测到的 bbox 进行关联的标记为 lost 状态。
- 7) 对于 lost 状态目标, id 记为 -1, 先从过去的 K_{lost} 帧找出可能丢失的目标 T_{pre} , 并根据(6)式计算相似性距离,如果小于阈值 K_{same} 则认为是目标重现,即为图3的 a_6 路径,否则标记为新的目标,赋予一个新的目标序号。
- 8) 根据当前帧继续训练新的 tracker,重复步骤3) ~ 7),直到序列结束。

4 实验

首先介绍了多目标追踪的评价指标;然后通过实验一分析不同参数对本文算法追踪效果的影响,选出最佳参数组合;通过实验二比较 IOU-tracker、SORT 和本文算法的性能;通过实验三分析本文算

法与其他优秀算法之间的对比。

本文算法运行在 Win10 操作系统(64 位), CPU 型号为 i7-4710MQ 8 核, 内存为 8G。先利用 VS2013 运行 Faster-RCNN 网络对图像序列进行目标检测, 将结果保存到文件中; 然后利用 matlab2016 编写 KCF_IOU 代码进行多目标追踪。

4.1 评价指标

以下评价准则^[1]用于评估各个算法的性能, 箭头向上表示数据越大跟踪性能越好, 向下表示数据越小跟踪性能越好, 包括: 每帧误报警率(FAF↓)、正确跟踪轨迹占 80% 以上的百分比(MT↑)、正确跟踪的轨迹占 20% 以上的百分比(ML↓)、目标 ID 的改变次数(IDS↓)、误检目标的数目(FP↓)、漏检目标的数目(FN↓)、轨迹断开的次数(FM↓)、运行速度(Hz↑)、多目标跟踪准确度(MOTA↑)和多目标跟踪精度(MOTP↑)。

多目标跟踪准确度(MOTA↑), 计算方法如下:

$$M_{\text{accuracy}} = 1 - \frac{\sum_t (n_t + p_t + s_t)}{\sum_t g_t}, \quad (7)$$

式中: t 表示帧序; n_t 、 p_t 、 s_t 分别为指标 FN、FP、IDS; g_t 为真实目标数量; 多目标跟踪精度(MOTP↑), 计算方法如下:

$$M_{\text{precision}} = \frac{\sum_{t,i} d_{t,i}}{\sum_t c_t}, \quad (8)$$

式中: c_t 表示第 t 帧正确追踪的结果数量; $d_{t,i}$ 表示第 i 个目标在第 t 帧追踪结果的 bbox 与真实 bbox 重叠率。

4.2 实验一: 探讨不同参数对实验的影响

在 KCF_IOU 追踪算法中, 主要影响变量参数有 3 个, 分别为 K_{lost} 、 K_{same} 、 K_{score} 。因此, 对这 3 个参数分别设置一些候选参数, 希望通过对候选参数进行组合, 训练得出一组较好的组合, 参数取值范围见表 1。

经过 3 个变量的组合, 使用 2DMOT2015 训练集运行 64 次后得到对应参数的实验结果, 如图 5 所示。图 5 的横坐标为不同的 K_{lost} 取值, 纵坐标分别为不同的 K_{same} 取值, 每一层为不同的 K_{score} 取值, 层

表 1 算法可选参数范围

Table 1 Selectable parameter range of algorithm

Parameter	K_{lost}	K_{same}	K_{score}
Range	2, 4, 6, 8	10, 20, 30, 40	0.6, 0.7, 0.8, 0.9

中线与线的交点表示不同评价指标值。

当把检测器行人置信度 K_{score} 从 0.6 改变到 0.9 时, 图 5(a) 的 MOTA 值从 31.555 提高到 38.166, MOTA 指标明显变好, 分析认为, 把置信度逐步提高, 去掉了非行人的情况, 使得算法能够在少量目标的情况下拥有更好的准确性能。相应的图 5(b) 的 MOTP 值也从 71.526 提高到 72.452, MOTP 指标也在逐渐提高, 但是每一层的值是相同的, 分析(8)式可知, MOTP 值是衡量目标位置误差的指标, 而位置误差和检测器密切相关, 由于本实验用的都是同一个检测器, 因此该指标不变。

图 5(h) 是 IDS 的实验结果, 由图可知, 置信度逐渐提高、目标相似阈值逐渐增加、最大丢失时间逐渐延长, 可以较好地减少 IDS 数量。根据(6)式可知, 目标相似阈值变大时, 允许目标有更好的稳健性。当最大丢失时间从 2 帧扩大到 8 帧时, IDS 数量从 757 降低到 670。理论上, 若有一个更好的目标相似性判定方法, 然后把 K_{lost} 设置得足够大, 就能找回丢失的目标, 但也会带来处理耗时的问题, 所以仅选择最大生存时间为 8 帧, 如果目标丢失时间超过 8 帧则认为目标彻底消失。

观察图 5(c)~(g) 和 (i) 等指标, 发现此类指标与 K_{same} 、 K_{lost} 取值无关, 仅随检测器置信度的增加而增加。分析认为, 这些评价指标都是根据目标位置进行的判定, 而算法在初始化时直接使用检测器检测到的目标位置, 没有在算法内部进行修正, 因此评价指标不变。

根据本实验, 算法最终采用参数为 $K_{\text{lost}} = 8$, $K_{\text{same}} = 40$, $K_{\text{score}} = 0.9$ 的组合。

4.3 实验二: 快速处理方法之间的对比

IOU-tracker 算法^[7]和 SORT 算法^[8]都是基于 IOU, 因此具有快速处理的优势, 不同之处在于, SORT 算法添加了卡尔曼滤波^[15]进行位置预测, 于是将这两个方法和本文算法进行对比。本实验采用 2DMOT2015 数据集的训练数据, 实验结果如图 6 所示, 横坐标为不同指标, 纵坐标是将不同指标通过缩放方式限定在 0~80 之间, 目的是便于直观比较各个指标间的差距。在算法处理速度方面, IOU 的处理速度可达到 100000 frame/s, SORT 处理速度达到 260 frame/s, 本文算法仅为 30 frame/s, 这是因为本文算法使用了外观模型, 需要对图像序列进行一些额外计算, 从而导致速度受到了影响。在算法处理准确性方面, 本文算法的 MOTA 指标高出 SORT 算法 10.3%; FP 指标极大地优化, 因为算法

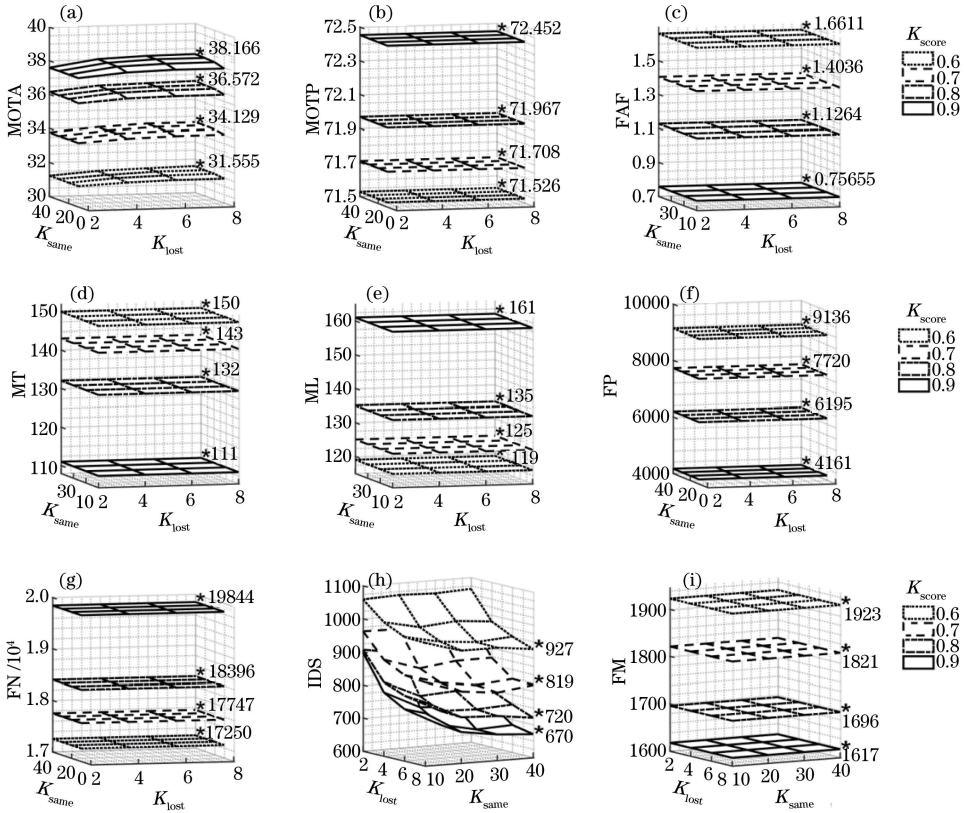


图5 不同参数的影响。(a) MOTA;(b) MOTP;(c) FAF;(d) MT;(e) ML;(f) FP;(g) FN;(h) IDS;(i) FM
Fig. 5 Effects of different parameters. (a) MOTA; (b) MOTP; (c) FAF; (d) MT; (e) ML; (f) FP; (g) FN; (h) IDS; (i) FM

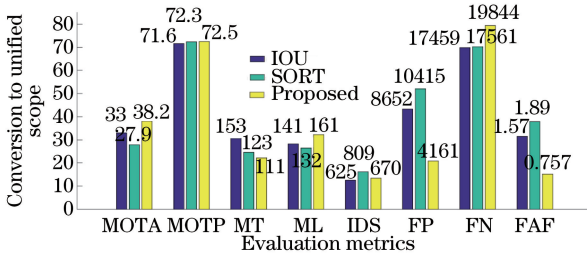


图6 IOU、SORT 和本文算法的对比
Fig. 6 Comparison among IOU, SORT, and proposed algorithm

加入对外观模型的考虑,提高了判别能力,有效减少了误检情况;MOTP 指标变化不大,这是因为本实验均采用相同的检测器;FAF 指标相对其他两个算法也有明显优化。综合分析,本文算法通过牺牲处理速度,实现了准确性方面的提升。

4.4 实验三:与优秀算法比较

为了更好地证明本文算法的追踪性能,使用 2DMOT2015 和 MOT16 测试集,和其他算法进行指标比较。MOT16 数据集相比于 2DMOT2015,权衡了训练集和测试集之间的相关性、对 bbox 标注做了更精准的改进。

2DMOT2015 数据集的实验数据见表 2,从表 2 可以看出,本文算法在 MOTA 指标上能够达到 29.2%,并且可以保证处理速度达到 30 frame/s。对比 DP_NMS 算法^[16],其处理速度比本文算法快 15 倍,但准确性比本文算法低 15%;对比 LDCT 算法^[17],该方法在多个指标上均表现出很好的效果,但是总体效果 MOTA 的值仅为 4.7%,分析原因发现该算法仅支持相机静止情况,未考虑相机运动情况;对比 TC_ODAL 算法^[18],IDS 数量仅为 637,但是其他指标却不好,这是由于该算法使用了更优秀的外观模型进行追踪;对比 CEM 算法^[5],FM 指标仅为本文算法的 1/3,大幅减少了轨迹断裂的情况,这是因为 CEM 算法是离线处理过程,通过对追踪结果不断优化,有效地解决了轨迹断开的问题,但是带来严重的误检。综合表 2 数据来看,本文算法具有在线追踪的能力,不仅具有较高的准确性,而且具有较快的处理速度。

MOT16 数据集的实验数据见表 3,由于此数据集的平均密集程度,从 2DMOT2015 的 10.6 提升到 30.8,所以 CEM、DP_NMS 算法的处理速度大幅降

表2 本文方法在2DMOT2015序列上的表现

Table 2 Performance of proposed approach on 2DMOT2015 sequences

Tracker	MOTA	MOTP	FAF	MT / %	ML / %	FP	FN	IDS	FM	Hz
DP-NMS	14.5	70.8	2.3	6.0	40.8	13171	34814	4537	3090	444.8
LDCT	4.7	71.7	2.4	11.4	32.5	14066	32156	12348	2918	20.7
TC_ODAL	15.1	70.5	2.2	3.2	55.9	12970	38538	637	1716	1.7
CEM	19.3	70.7	2.5	8.5	46.5	14180	34591	813	1023	1.1
Proposed	29.2	71.1	0.9	10.0	34.5	5369	36590	1513	2859	30.0

低。EAGS16算法^[19]通过对目标进行超像素分割建立外观模型并追踪,在目标遮挡和拥挤场合得到很好的效果,因此该算法的IDS仅为575,远小于本文算法,并且拥有较快的处理速度,但存在误检和漏检问题。SMOT算法^[20]为了适应目标非惯性问题,利用广义线性分配方法构建运动模型,较好地解决了相机运动等问题,但是由于外观模型

不理想,导致追踪过程不连续,FM值较高。MAP-RF算法^[6]的FP指标比本文算法高10%,因为该算法构建了多个外观模型来实现位置预测,但是存在误检情况,这是由于外观模型过拟合导致目标追踪出现偏差。本实验中,算法依旧表现出55.8%的准确性,明显高于其他算法,误检和漏检情况也优于其他算法。

表3 本文方法在MOT16序列上的表现

Table 3 Performance of proposed approach on MOT16 sequences

Tracker	MOTA	MOTP	FAF	MT / %	ML / %	FP	FN	IDS	FM	Hz
DP-NMS	26.2	76.3	0.6	4.1	67.5	3689	130557	365	638	5.9
CEM	33.2	75.8	1.2	7.8	54.4	6837	114322	642	731	0.3
EAGS16	47.4	75.9	1.4	17.3	42.7	8369	86931	575	913	197.3
SMOT	29.7	75.2	2.9	5.3	47.7	17426	107552	3108	4483	0.2
MAP-RF	64.8	79.7	—	37.8	22.3	7780	57081	645	—	30.0
Proposed	55.8	76.5	1.2	22.8	33.3	7046	71601	1856	3114	15.0

5 结 论

对于多目标追踪问题,可以分为两个部分:构建目标追踪的外观模型,利用图像色彩、角点、轮廓等信息进行追踪;构建目标追踪的运动模型,利用位置信息进行追踪。然后将两部分结合起来实现多目标追踪过程。很多时候算法的瓶颈在于外观模型的计算,有些场景需要有较快的处理速度,如果仅利用目标位置进行追踪,会导致目标追踪的准确性有很大缺陷,然而在考虑利用目标的图像信息进行处理时,图像的处理速度又会大幅减慢。因此,本文利用KCF算法充当多目标追踪的外观模型,以提高追踪的准确性。通过实验二的结果可以分析出,加入外观模型会把处理速度降低为30 frame/s,但追踪的准确性可有效提高10%以上。实验三中使用2DMOT2015、MOT16数据集,和其他优秀算法进行对比,当采用基于HOG特征的KCF算法时,综合结果优于其他算法。本文方法在追踪速度和追踪准确性方面可以到达一个很好的权衡。此外,由于处理过程仅涉及历史帧,属于在线处理算法,可以将该算法应用在需要实时处理的场景,满足在线处理

数据的要求。

为了保持较快的处理速度,对于目标相似性判断仅使用简单的灰度模板比较方式。虽然能够减少一定的计算量,但效果却不是很好。对于目标重现判定要求严格的场景,可以采用提取目标HOG特征,然后计算余弦距离等。

由于KCF算法没有考虑目标尺度变化问题,因此本文算法训练的KCF追踪器仅使用前后连续两帧进行训练和预测,后续将考虑多尺度情况,把目标追踪器的训练过程加长,以保持更好的稳定性。

参 考 文 献

- [1] Leal-Taixé L, Milan A, Reid I, *et al.* MOTChallenge 2015: towards a benchmark for multi-target tracking [EB/OL]. (2018-04-08) [2018-11-25]. <https://arxiv.org/abs/1504.01942>.
- [2] Kim C, Li F X, Ciptadi A, *et al.* Multiple hypothesis tracking revisited [C] // 2015 IEEE International Conference on Computer Vision (ICCV), December 7-13, 2015, Santiago, Chile. New York: IEEE, 2015: 4696-4704.

- [3] Xiang Y, Alahi A, Savarese S. Learning to track: online multi-object tracking by decision making[C]// 2015 IEEE International Conference on Computer Vision (ICCV), December 7-13, 2015, Santiago, Chile. New York: IEEE, 2015: 4705-4713.
- [4] Chen L, Ai H Z, Shang C, *et al.* Online multi-object tracking with convolutional neural networks[C]// 2017 IEEE International Conference on Image Processing (ICIP), September 17-20, 2017, Beijing, China. New York: IEEE, 2017: 645-649.
- [5] Milan A, Roth S, Schindler K. Continuous energy minimization for multitarget tracking [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014, 36(1): 58-72.
- [6] Zhou H Y, Yang Y, Wang S Y. Multiple object tracking algorithm based on kernel correlation filter [J]. Laser & Optoelectronics Progress, 2018, 55(9): 091502.
周海英, 杨阳, 王守义. 基于核相关滤波器的多目标跟踪算法[J]. 激光与光电子学进展, 2018, 55(9): 091502.
- [7] Bochinski E, Eiselein V, Sikora T. High-speed tracking-by-detection without using image information [C]// 2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), August 29-September 1, 2017, Lecce, Italy. New York: IEEE, 2017: 8078516.
- [8] Bewley A, Ge Z Y, Ott L, *et al.* Simple online and realtime tracking [C]// 2016 IEEE International Conference on Image Processing (ICIP), September 25-28, 2016, Phoenix, AZ, USA. New York: IEEE, 2016: 3464-3468.
- [9] Wojke N, Bewley A, Paulus D. Simple online and realtime tracking with a deep association metric[C]// 2017 IEEE International Conference on Image Processing (ICIP), September 17-20, 2017, Beijing, China. New York: IEEE, 2017: 3645-3649.
- [10] Henriques J F, Caseiro R, Martins P, *et al.* High-speed tracking with kernelized correlation filters[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(3): 583-596.
- [11] Dalal N, Triggs B. Histograms of oriented gradients for human detection [C]// 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), June 20-25, 2005, San Diego, CA, USA. New York: IEEE, 2005: 886-893.
- [12] Cheng Y, Li J Z, Zhu L N, *et al.* Correlation filter tracking algorithm based on model and scale updating [J]. Laser & Optoelectronics Progress, 2018, 55(12): 121015.
成悦, 李建增, 褚丽娜, 等. 基于模型与尺度更新的相关滤波跟踪算法[J]. 激光与光电子学进展, 2018, 55(12): 121015.
- [13] Ren S Q, He K M, Girshick R, *et al.* Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [14] Tian Q, Yuan T Y, Yang D, *et al.* A pedestrian detection method based on dark channel defogging and deep learning [J]. Laser & Optoelectronics Progress, 2018, 55(11): 111007.
田青, 袁瞳阳, 杨丹, 等. 基于暗通道去雾和深度学习的行人检测方法[J]. 激光与光电子学进展, 2018, 55(11): 111007.
- [15] Kalman R E. A new approach to linear filtering and prediction problems [J]. Journal of Basic Engineering, 1960, 82(1): 35-45.
- [16] Pirsivash H, Ramanan D, Fowlkes C C. Globally-optimal greedy algorithms for tracking a variable number of objects [C]// CVPR 2011, June 20-25, 2011, Colorado Springs, CO, USA. New York: IEEE, 2011: 1201-1208.
- [17] Solera F, Calderara S, Cucchiara R. Learning to divide and conquer for online multi-target tracking[C]//2015 IEEE International Conference on Computer Vision (ICCV), December 7-13, 2015, Santiago, Chile. New York: IEEE, 2015: 4373-4381.
- [18] Bae S H, Yoon K J. Robust online multi-object tracking based on tracklet confidence and online discriminative appearance learning [C]// 2014 IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2014, Columbus, OH, USA. New York: IEEE, 2014: 1218-1225.
- [19] Sheng H, Zhang X Y, Zhang Y, *et al.* Enhanced association with supervoxels in multiple hypothesis tracking[J]. IEEE Access, 2019, 7: 2107-2117.
- [20] Diele C, Camps O I, Sznaier M. The way they move: tracking multiple targets with similar appearance[C]//2013 IEEE International Conference on Computer Vision, December 1-8, 2013, Sydney, NSW, Australia. New York: IEEE, 2013: 2304-2311.