

doi:10.11835/j.issn.1000-582X.2017.07.007

基于图的流行排序的视觉跟踪

邱慧丽^{1a,1b}, 宋启祥^{1a,1b}, 赵楠²

(1.宿州学院 a.智能信息处理实验室;b.信息工程学院,安徽宿州 234000;

2.安徽大学 计算机科学与技术学院,安徽合肥 230601)

摘要:为了克服基于检测的目标跟踪中的模型漂移问题,在基于检测的目标跟踪框架下提出一种新的基于图的流行排序的目标跟踪方法。该方法能够抑制在跟踪过程中目标变形、尺度变化以及遮挡等带来的背景信息的影响。首先,把目标矩形框划分为不重叠的图像块,构造一个k-正则图,即以这些图像块为图结点,构造k-正则图,边权定义为结点之间的底层特征的相似性。其次,为每一个图像块分配一个权重,用于表示该图像块在目标表达中的重要性,以此抑制背景信息的干扰,并通过半监督的方式进行计算。特别初始化一些背景或目标图像块,设其权重为1,其他为0,通过流行排序算法计算所有图像块属于背景或目标的权重。此外,使用多尺度特征金字塔的方法处理跟踪过程中的目标尺度变化,同时提高了初始图像块的可靠性。最后,把所有图像块的底层特征进行加权,连接成一个向量,作为矩形框的特征表达,并使用结构化输出(Struck)算法进行跟踪。在几个公共视频序列上进行了实验,结果表明,研究方法的跟踪性能极大地超过了其他跟踪方法。

关键词:视觉跟踪;流行排序;尺度处理;Struck 算法

中图分类号:TP 391.41

文献标志码:A

文章编号:1000-582X(2017)07-043-09

Graph-based manifold ranking for visual tracking

QIU Huili^{1a,1b}, SONG Qixiang^{1a,1b}, ZHAO Nan²

(1a. Intelligent Information Processing Lab; 1b. School of Information Engineering, Suzhou University, Anhui 234000, P.R.China; 2. School of Computer Science and Technology, Anhui University,

Hefei 230601, Anhui, P.R.China)

Abstract: This paper proposes a novel object tracking approach via graph-based manifold ranking to handle the model drift problem in the tracking-by-detection framework. The proposed approach can suppress the effects of background information caused by object deformation, scale variation and occlusion in object tracking. First, we partition the target bounding box into non-overlapping image patches, and take these

收稿日期:2017-02-10

基金项目:国家自然科学基金资助项目(61374128,41173106,41373095);安徽省科技攻关计划(1501zc04048);安徽省教育厅自然科学研究产学研重点项目(KJ2014A247);宿州学院智能信息处理实验室开放课题资助(2016ykf13)。

Supported by National Natural Science Foundation of China(61374128,41173106,41373095); Anhui Scientific and technological project(1501ZCD4048); Natual Science Industry University Research Project of Anhui Provincial Education Department(KJ2014A247); Openjing Foundation of Intelligent Information Processing Laboratory,Suzhou University(2016ykf13).

作者简介:邱慧丽(1989-),女,博士,主要从事图像处理、模式识别、环境水文地球化学等方向的研究,
(Tel)18905578734;(E-mail)qiu_hui_li@126.com。

image patches as graph nodes to construct k -regular graph, in which the edge weight between two neighbor nodes are measured by the distance of their low-level features. Second, we assign each patch with a weight describing the importance in representing the object, and compute it in a semi-supervised way. In particular, we initialize some patches as object patches with the weights 1, and remaining patches with the weights 0. The graph-based manifold ranking is then performed to obtain the weights of all patches. Moreover, we propose to determine the optimal scale based on multi-scale feature pyramid to address scale adaptation while improving the quality of initial patches in object tracking. Finally, we concatenate all weighted patch descriptors into a vector to represent the bounding box feature, and then integrate it into structure output (Struck) algorithm to carry out object tracking. Experimental results on several public video sequences suggest that the proposed method significantly outperforms other tracking methods.

Keywords: visual tracking; manifold ranking; scale adaptation; Struck algorithm

视觉跟踪(又称单目标跟踪)是在给定初始状态(第一帧的目标矩形框)的条件下,在后续帧中持续稳健地跟踪目标物体。它是目前图像处理、模式识别和计算机视觉中的一个经典问题和重要研究内容,也是智能视频系统的关键和核心,有着非常广泛的应用前景,如自动驾驶、监控系统和现实增强等。因此,越来越多的学者开始投入到视觉跟踪的研究中,并且已经取得了丰硕的成果^[1-12]。由于在跟踪过程中可能出现目标突然运动、周围场景变化、遮挡等情况,很难实现高效稳健地目标跟踪,目前,解决跟踪问题主要分为基于检测^[1-2]和基于生成^[3]2种跟踪算法,近来,基于检测的目标跟踪算法因其较好的跟踪效果,得到了众多学者的青睐。

基于检测的目标跟踪算法的大致过程如下:给定目标的初始状态(第一帧的目标矩形框),分别使用正、负样本训练一个目标检测器。在后续的视频帧中,用该检测器对跟踪位置附近的样本进行检测打分,得分最高的样本作为跟踪结果。同时,根据跟踪结果选择更新分类器。但是,实际情况中,当跟踪结果含有一些背景等干扰目标的信息时,更新后的检测器的判别能力就会慢慢的降低,逐渐偏离真实的目標模型,直至导致跟踪失败,即模型漂移问题。总结发现,在跟踪过程中跟踪结果(目标矩形框)包含背景信息的主要原因包括:1)矩形框无法准确描述不规则形状的目标;2)跟踪过程中目标尺度的变化;3)部分遮挡和完全遮挡的影响。如图 1 中的(a) (b) (c)所示。

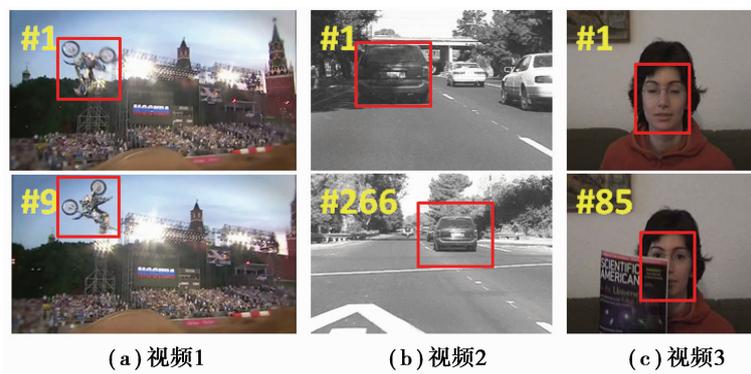


图 1 3 个挑战视频的示例

Fig.1 Three typical samples of challenging videos

针对以上问题,许多学者致力于消除分类器更新过程中的背景影响。其中,文献[11-16]根据与目标矩形框中心点的距离来简单的对像素点加权,文献[17-19]在跟踪模型中引入分割结果,以此消除背景的干扰。但前者忽略了目标的真实形状,而后者依赖于分割结果,分割结果的好坏影响很大,具有较差的通用性。最近也有许多学者使用深度学习训练神经网络用于跟踪^[20-21],同样取得很好的跟踪效果,但深度学习对机器的运算能力和性能有很高的要求,其模型和运算复杂度也超过传统跟踪方法。

最近, Kim 等^[18]提出一种基于空间排序加权图像块的矩形框特征描述方法,能够较好的表示物体的特

征,以此抑制背景信息在目标表达中的干扰,并且取得了非常好的跟踪性能。该方法初始化一部分目标和背景图像块作为先验种子点。利用图像块之间的相似性关系把初始种子点的权值通过改进的随机游走算法传播到其他图像块上,以此得到所有图像块的权重;把求得的权重嵌入到矩形框目标特征中,实现包含目标的图像块对应权重大,而包含背景的图像块权重相应较低,从而达到削弱背景的干扰和影响,然后结合 Struck 算法^[16]进行跟踪。该方法主要具有以下2点不足:1)该随机游走算法是一个迭代算法,具有一定的计算复杂度;2)没有对目标做尺度处理,会导致种子点选取错误或不合理,直接造成目标或背景图像块的权重初始化不准确,无法有效抑制背景信息的影响,极大地限制了该方法的性能。

为此,提出一种基于图的流行排序的目标跟踪方法,能够在复杂场景和环境下稳健并高效地跟踪目标。一方面,使用流行排序算法计算每个图像块的权重,该算法具有闭合解,因此,具有较高的求解效率。另一方面,根据给定当前帧的跟踪位置,在下一帧的这个位置周围设置一个搜索区域,在搜索区域中建立多尺度特征金字塔^[22-23],解决跟踪过程中尺度变化带来的影响。把计算的权重和尺度处理融合到 Struck 算法^[16]中,实现目标的高效稳健跟踪。实验表明,提出的方法极大地提高了跟踪性能。

1 算法

研究提出的算法利用流行排序算法获得图像块加权表示目标信息来减少实际跟踪过程中背景干扰的问题,首先介绍加权空间图像块特征描述和 Struck 跟踪框架的概念。

1.1 基于加权空间排序图像块的特征描述

视觉跟踪过程中,一般是通过一个矩形框来定位表示跟踪的目标,而矩形框内的特征表达对跟踪性能往往具有较大的影响。一种较为鲁棒的特征描述方法就是基于加权空间排序图像块的特征描述,这种描述方法被广泛地应用于计算机视觉中的很多技术上,如梯度方向直方图(HOG),HOG 特征表示的是把一个矩形框中的所有局部梯度方向直方图连接在一起作为该矩形框的特征表达,这种特征描述成功地应用于很多的视觉问题上,包括目标检测和跟踪等^[24-26]。研究则使用这种特征描述来表达一个矩形框,并且每个局部图像块可以使用多种特征进行表达。

假设在第 t 帧,为了对矩形框 y 进行可靠地表达,一种做法是把矩形框划分成 n 互不重叠的局部图像块,每个图像块 i 提取一系列底层特征 f_i 进行表示。然后,把所有图像块的特征按图像块的空间顺序连接在一起,就得到了基于加权的空间排序图像块的特征描述

$$\Phi(x_t, y) = [f_1^T, \dots, f_n^T]^T, \quad (1)$$

这种特征通过保持局部图像块特征的空间顺序能够很好地表达矩形框内的结构信息,

1.2 Struck 跟踪算法

在第 t 帧 x_t , 结构化输出(Struck)算法是通过最大化分类得分 $\langle w, \Phi(x_t, y) \rangle$ 来估计最优的目标矩形框的位置 y_t

$$y_t = \arg \max_y \langle w, \Phi(x_t, y) \rangle. \quad (2)$$

其中, w 是第 $t-1$ 帧的决策平面的法向量。在更新分类器时, Struck 能够利用结构化的样本消除样本标签的歧义性。这里,结构化样本是指同一帧中的一个目标矩形框和附近的矩形框组成的。Struck 则约束目标矩形框的分类分数要大于附近的矩形框一个边界(margin),而该边界由这2个矩形框的重合比决定的。因此, Struck 不需要一些具有二值标签的样本进行训练分类器,进而避免了分类器更新过程中的标签歧义性问题。Struck 表现出了优秀的跟踪性能,其训练学习策略也被最近的跟踪方法^[27-30]所采用。

2 研究方法

类似于文献^[17],通过把目标矩形框划分成多个局部空间有序的图像块,并且为每个图像块分配一个表达目标的权重,通过权重可以削弱目标跟踪中背景信息的干扰。

2.1 图模型构建

由于空间邻接的图像块具有相似的外观和权重,因此,构建一个 2-正则图 $G = (V, E)$,其中, V 表示由图像块构成的结点集合, E 表示无向边集,每个结点只与空间相邻的 8 个结点相连一条边。为每个图像块提取 d 维梯度方向直方图和 RGB 颜色直方图作为特征表示,记为 f_i 为第 i 个图像块的特征,每 2 个结点之间的权重通过它们之间的特征距离来衡量

$$w_{i,j} = \exp(-\gamma \|f_i - f_j\|^2), \quad (3)$$

其中, r 是一个控制边缘强度的参数。可以看出, $w_{i,j}$ 越大,图像块 v_i 和 v_j 越相似,反之亦然。

2.2 基于流行排序算法的图像块权重计算

为了得到精确且有效的目标描述,将目标矩形框划分为 n 个无重叠的局部图像块,并通过公式(1)构造目标的空间排序图像块特征描述。但是由于目标的形状、尺度变化以及遮挡等因素,会导致某些图像块属于背景,这样,如果等同地对待所有图像块,则会极大地降低跟踪性能,最终导致模型漂移问题。为了解决这个问题,为每个图像块分配一个权重,用来表达对目标的贡献,即权重越大,则该图像块属于目标的可能性越大,反之属于背景的可能性越大。下面则通过流行排序算法计算这些图像块的权重。

1) **流行排序算法**。流行排序是一个半监督学习的方法,即对于一个图模型,给定一些种子结点,利用结点之间的内在流行结构对每个结点进行排序,得到每个结点的排序得分。具体地,给定结点集合 $V = \{v_1, v_2, \dots, v_n\} \in R^{d \times n}$,且选取其中一些结点为种子点,而其余的结点则为根据与种子点相似性关系待排序的结点。记 $g: V \rightarrow R^n$ 为一个排序函数,即为每一个结点 v_i 计算一个排序得分 g_i ,那么 g 可以看成是一个得分向量 $g = [g_1, g_2, \dots, g_n]$ 。记 $g_0 = [g_{0,1}, g_{0,2}, \dots, g_{0,n}]$ 为种子结点的指示函数,即如果 v_i 为种子点,则 $g_{0,i} = 1$,否则 $g_{0,i} = 0$ 。则最优的排序结果可以通过求解下列的优化函数

$$g^* = \arg \min_g \frac{1}{2} \left(\sum_{i,j=1}^n w_{ij} \left\| \frac{g_i}{\sqrt{d_{ii}}} - \frac{g_j}{\sqrt{d_{jj}}} \right\|^2 + \mu \sum_{i=1}^n \|g_i - g_{0,i}\|^2 \right). \quad (4)$$

其中, μ 表示权衡参数, $D = \text{diag}\{d_{11}, d_{22}, \dots, d_{nn}\}$ 为图 G 的边的亲和性矩阵 $W = [w_{ij}]_{n \times n}$ 的度矩阵。公式(4)的第一项和第二项分别为平滑约束和拟约束,即好的排序函数要使得邻近的、相似的结点之间的排序得分尽可能相近,同时,要使得排序得分和种子点的分数尽可能地接近。该公式的最优解可以通过把目标函数求导后等于 0 得到

$$g^* = (I - \alpha S)^{-1} g_0. \quad (5)$$

其中, $\alpha = \frac{1}{\mu + 1}$, $S = D^{-1/2} W D^{-1/2}$ 为归一化后的拉普拉斯矩阵, I 表示单位矩阵。



图 2 区域划分示意图

Fig.2 Illustration of region partition

2) **权重计算**。把一部分的图像块初始为目标种子点和背景种子点,图 2 示意了这一过程。图中,红色矩形框表示标注真值或跟踪结果,黄色和蓝色矩形框则为收缩和放大后的矩形框。把黄色矩形框内的图像块初始化为目标种子结点,把红色矩形框和蓝色矩形框之间的图像块初始化为背景种子结点,则其余的图像块则认为是不确定的。根据初始化的种子点类别不同,利用流行排序算法分别计算每个图像块的属于目标和背景的权重,记为 $\mathbf{g}^f = [g_1^f, g_2^f, \dots, g_n^f]$ 和 $\mathbf{g}^b = [g_1^b, g_2^b, \dots, g_n^b]$ 。由此可得到 t 时刻最终 2 种方式联合的图像块的权重 $\rho_i(t)$ 为

$$\rho_i(t) = \frac{1}{1 + \exp(-\beta(g_i^f(t) - g_i^b(t)))} \quad (6)$$

其中, β 控制逻辑函数的峭度。研究对矩形框中的对每个图像块的特征加权,即可得到基于空间排序加权图像块的特征描述为 $\Phi = [\rho_1(t)\mathbf{f}_1^T, \rho_2(t)\mathbf{f}_2^T, \dots, \rho_n(t)\mathbf{f}_n^T]$ 。该特征描述符能够有效地表达和区分出目标图像块和背景图像块,从而适当消除背景信息对目标的干扰。

2.3 目标跟踪

上述的目标描述仍有一个比较大的缺点,就是在尺度变化时,种子点往往会初始化错误,如图 3 所示。图中,第一行分别表示第一帧的标注真值、没有尺度处理和有尺度处理的矩形框情况,第二行分别为对应的计算的权重结果图。从图中可以看出,随着尺度的变化,如果不进行尺度处理,则红色矩形框无法较好的表示目标。这样,会导致种子点的错误初始化,进而使得权重的计算非常不准确,无法有效地克服背景的干扰。因此,提出一种基于多尺度特征金字塔的方法进行尺度处理,分为以下 3 步:

1) 权重计算。根据第一帧的结果或者后续帧(第 $t-1$ 帧 x_{t-1})的跟踪结果,使用基于图的流行排序算法计算每个划分图像块的权重。

2) 在第 t 帧,首先,根据第 $t-1$ 帧的跟踪结果,在其附近设置搜索区域,并建立多尺度特征金字塔。其次,对于任一尺度的样本,使用计算的权重得到该样本的加权空间排序图像块特征 $\Phi(x_t, y)$ 。最后,根据第 $t-1$ 帧训练的分类器,通过公式(2)最大化分类器得分来得到最优的候选样本以及该样本所对应的尺度,分别作为该帧的跟踪结果以及最优的尺度。

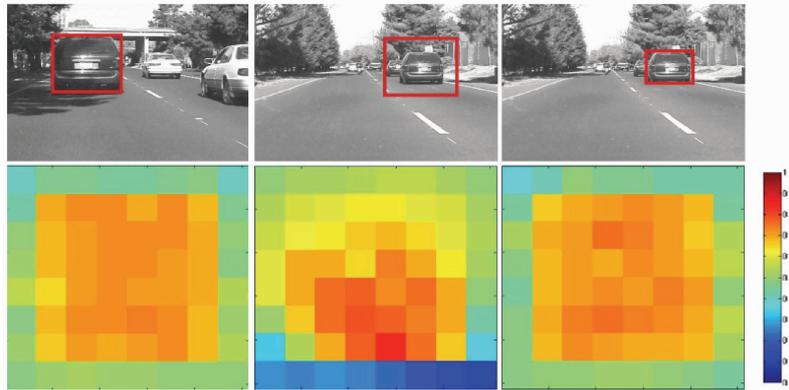


图 3 公共序列 Car4 上,自适应尺度在权重计算中的作用示例

Fig.3 Effectiveness of scale adaptation in computing weights on public video sequence Car4

3) 根据跟踪结果 y_t 及其特征 $\Phi(x_t, y)$,通过 Struck 算法^[15]更新该帧的分类器。特别地,为了克服目标外观突然形变、光照变化和遮挡等因素的影响,根据跟踪结果的可靠程度来决定是否更新分类器,以此避免分类器受噪声的影响而导致模型漂移。具体地,定义置信度得分 θ_t 来衡量在第 t 帧的跟踪结果置信度

$$\theta_t = \frac{1}{|s_t|} \sum_{s \in s_t} \langle S, \Phi(x_t, y_t) \rangle. \quad (7)$$

其中, S_t 是时间间隔 t 内的正支持向量的集合, 只有当置信度得分大于某一预设阈值 η 时才会更新分类器。

3 实验部分

3.1 实验设置

1) 数据集。在标准数据集中选取了具有代表性的 6 个具有较大挑战的公共视频序列用来评价提出的跟踪方法。

2) 参数设置。实验中所有设置都固定, 计算相邻图像块特征距离时 $r=5.0$, 计算联合图像块权重 $\beta=35.0$, 结构化输出 SVM 使用线性核函数, 置信度得分阈值 $\eta=0.3$, 即当得分满足大于阈值条件时, 分类器要在当前帧进行更新操作。为了提高计算效率, 将每一帧的跟踪矩形框尺寸进行缩放变换, 使得矩形框的最小边长是 32 像素值。搜索窗口的边长被固定为 $2\sqrt{WH}$, W 和 H 分别对应缩放变化处理之后的跟踪矩形框的宽和高。

3) 对比方法。为了说明算法的有效性与先进性, 与 6 种当前流行的跟踪方法做对比, 包括 Struck^[16], TLD^[17], ASLA^[22], SCM^[23], LSHT^[19] 和 DSST^[24]。

4) 度量标准。采用了被广泛应用的衡量跟踪性能的标准成功率, 跟踪精度得分来评价所有的跟踪方法。其中 \cap 和 \cup 分别代表 2 个区域的交和并, $|\cdot|$ 表示区域中像素的个数, 预设阈值固定为 0.5。跟踪精度分数是指跟踪结果矩形框与标注真值矩形框中心位置平均欧式距离小于预设阈值得到的, 其中, 跟踪成功帧数比例是指跟踪结果矩形框中心位置与标注真值矩形框中心位置小于阈值的帧数占总帧数的比例, 以此, 用来评估跟踪方法的整体跟踪性能, 预设阈值固定为 20。

3.2 对比结果与分析

研究与其它跟踪方法在 6 个公共视频序列上的对比结果如表 1 所示, 其中 6 个公共视频序列来自于 Wu Y^[31] 等人在 CVPR2013 上发表的标准数据集。从表中可以看出, 方法在整体上取得了最优的性能, 平均精度达到 97.35%, 且大幅度的超过了次优的跟踪方法 81.57%, 比次优算法提高了 19.34% 的精度, 验证了方法的有效性。具体地, 尽管方法在 Shaking 上比 DSST 稍差, 但是也取得了 99.20% 的精度。注意到, DSST 算法是在 VOT2014 视觉跟踪竞赛上取得了第一名。因为使用了经典的基于图的流行排序算法再结合尺度处理操作, 使得计算图像块前景概率时, 种子点选取更加合理, 计算结果更准确, 从而使得在其它 5 个序列均得到了最好的结果, 特别地, 是在 skiing 序列上, 由于该序列的目标较小、外观变换大、背景干扰严重,

表 1 每个视频序列上的重叠精度百分比, 最高和次高的结果分别用红色和蓝色字体表示

Table 1 Overlap ratio of each video sequence, where the red and blue fonts indicate the best and second best performance, respectively

算法	boy	Car4	CarScale	shaking	skiing	walking2	Average
Ours	100.00	100.00	84.90	99.20	100.00	100.00	97.35
DSST ^[24]	100.00	100.00	84.50	100.00	4.94	100.00	81.57
ASLA ^[22]	43.50	100.00	71.00	23.30	12.30	39.80	48.32
SCM ^[23]	43.90	37.50	66.70	55.60	9.88	100.00	52.26
Struck ^[16]	97.50	39.90	43.30	52.90	4.94	43.00	46.92
LSHT ^[19]	50.70	27.60	44.80	69.90	3.70	38.40	39.18
TLD ^[17]	82.90	24.00	68.70	3.29	7.41	20.80	34.52

且运动速度较快,因此给跟踪方法带来了极大的挑战性,也使得大部分跟踪方法跟踪失败,而研究在这个序列上获得了 100.00% 的跟踪精度;在 shaking 序列上,目标存在光照突变、旋转、形变等特征影响,也使得大部分跟踪方法跟踪的效果并不理想,研究确能取得 99.2% 的跟踪精度;另外,在 boy、Car4、CarScale 和 Walking 序列上,存在目标尺度发生明显变化、背景光照变化以及目标局部遮挡等挑战,方法的跟踪效果仍能超过其它方法,并取得非常高的跟踪精度。此外,图 4 展示了在 6 个视频序列上的典型帧的跟踪结果。



图 4 研究方法在 6 个公共数据集上的结果展示

Fig 4 Sample results of the proposed method on 6 video sequences

然而,方法在某些复杂或者非常有挑战的场景下,例如:目标及其周围背景发生严重的遮挡、光照变化以及旋转等挑战时,获得的跟踪效果依然不理想,如图 5 所示。

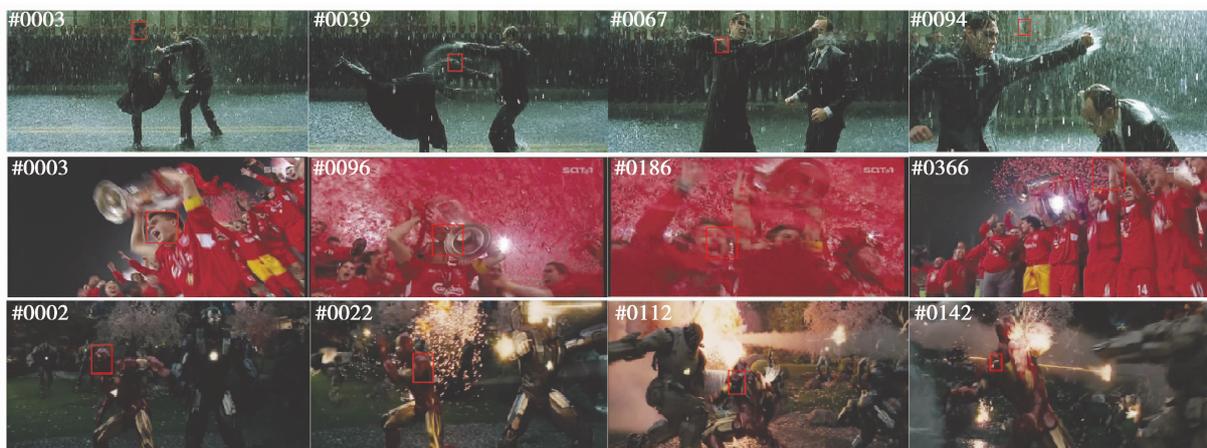


图 5 方法在复杂和具有挑战场景下的结果展示

Fig.5 Sample results of the proposed method on the challenging scenarios

4 结 语

提出了一种基于图的流行排序的鲁棒目标跟踪方法,利用空间加权有序图像块描述符和判别式分类器 SVM 来有效估计目标在后续帧中的位置以及目标真实尺度的变化。在 6 个挑战性的公共视频序列上与 6 种流行的目标跟踪方法进行了对比与分析,验证了方法的有效性。在未来的工作中,将从以下 3 个方面完善目前的工作,更为有效地抑制目标跟踪过程中的背景干扰:1)研究目标和背景图像块的初始化问题。因为简单地根据矩形框的位置初始化一些背景块,具有较差的通用性,因此需要根据不同的情形进行自适应地初始化;2)构造新的图模型。传统的 8-邻域图只考虑了图像块之间的空间局部邻接关系,因此考虑全局和局部相结合的方式建立新的图模型,更为有效地实现图像块之间的权重传播;3)通过引入一些先验知识研究更为鲁棒的基于图的流行排序算法。

参考文献:

- [1] Bai Y C, Tang M. Robust tracking via weakly supervised ranking svm[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Rhode Island, USA: IEEE,2012:1854-1861.
- [2] Yin Y J, Xu D, Wang X G, et al. Online state-based structured SVM combined with Incremental PCA for robust visual tracking[J]. IEEE Transactions on Cybernetics,2015,45:1988-2000.
- [3] Henriques J F, Rui C, Pedro M, et al. High-Speed tracking with kernelized correlation filters[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2015,37(3):583-596.
- [4] Avidan S. Ensemble tracking [J]. IEEE transactions on pattern analysis and machine intelligence,2008,29(2): 261-271.
- [5] Grabner H, Leistner C, Bischof H. Semi-supervised online boosting for robust tracking[C]// Proceedings of European Conference on Computer Vision. Marseille, France: IEEE,2008:234-247.
- [6] Mei X, Ling H B. Robust visual tracking using l1 minimization[C]// Proceedings of IEEE International Conference on Computer Vision. Kyoto, Japan: IEEE,2009:1436-1443.
- [7] Zhang T Z, Ghanem B, Liu S, et al. Robust visual tracking via multi-task sparse learning[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Rhode Island, USA: IEEE,2012,157(10):2042-2049.
- [8] Bao C L, Wu Y, Ling H B, et al. Real time robust l1 tracker using accelerated proximal gradient approach[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Rhode Island, USA: IEEE,2012,157(10): 1830-1837.
- [9] 侯跃恩,李伟光,四库,等.基于排名的结构稀疏表示目标跟踪算法[J].华南理工大学学报:自然版,2013,41(11):23-29.
HOU Yuen, LI Weiguang, SI Ku, et al. Target tracking algorithm with structured sparse representation based on ranks [J]. Journal of South China University of Technology Science and Technology,2013,41(11):23-29.(in Chinese).
- [10] 尹宏鹏,柴毅,魏洪波,等.Mean-shift 改进算法在火箭目标跟踪中的应用[J].重庆大学学报,2010,33(4):120-126.
YIN Hongpeng, CHAI Yi, WEI Hongbo, et al. Application of improved mean-shift algorithm in rocket target tracking[J].Journal of Chongqing University,2010,33(4):120-126.(in Chinese)
- [11] Zhang J M, Ma S G, Sclaroff S. MEEM: robust tracking via multiple experts using entropy minimization [C] // Proceedings of European Conference on Computer Vision. Zurich: IEEE,2014:188-203.
- [12] Babenko B, Yang M H, Belongie S, et al. Robust object tracking with online multiple instance learning [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2011,33(7):1619-1632.
- [13] 刘京诚,廖智勇,朱木健,等.一种基于服务机器人的视觉系统设计[J].重庆大学学报,2006,29(11):31-35.
LIU Jingcheng, LIAO Zhiyong, ZHU Mujian, et al. Design of vision system based on service robot[J].Journal of Chongqing University,2006,29(11):31-35.(in Chinese)
- [14] Lee D Y, Sim J Y, Kim C S. Visual tracking using pertinent patch selection and masking[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Columbus, OH, USA:IEEE,2014:3486-3493.

- [15] Comaniciu D, Ramesh D, Meer P. Kernel-based object tracking[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(5): 564-577.
- [16] Hare S, Saffari A, Torr P H S. Struck: Structured output tracking with kernels[C]// Proceedings of IEEE Conference on International Conference on Computer Vision. Rhode Island, USA: IEEE, 2012, 23(5): 263-270.
- [17] Kalal Z, Mikolajczyk K, Matas J. Tracking-learning-detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 34(7): 1409-1422.
- [18] Kim H U, Lee D Y, Sim J Y, et al. SOWP: Spatially ordered and weighted patch descriptor for visual tracking[C]// Proceedings of IEEE Conference on International Conference on Computer Vision. SANTIAGO, CHILE: IEEE, 2015: 3011-3019.
- [19] He S F, Yang Q X, Lau R W H, et al. Visual tracking via locality sensitive histograms[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Portland, OR, USA: IEEE, 2013, 9(4): 2427-2434.
- [20] Hu J L, Lu J W, Tan Y P. Deep metric learning for visual tracking[J]. IEEE Transactions on Circuits and Systems for Video Technology (T-CSVT), 2016, 26(11): 2056-2068.
- [21] Qi Y K, Zhang S P, Qin L, et al. Hedged deep tracking[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, Nevada, USA: IEEE, 2016: 4303-4311.
- [22] Jia X, Lu H C, Yang M H. Visual tracking via adaptive structural local sparse appearance model[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Rhode Island, USA: IEEE, 2012, 157(10): 1822-1829.
- [23] Zhong W, Lu H C, Yang M H. Robust object tracking via sparsity based collaborative model[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Rhode Island, USA: IEEE, 2012, 157(10): 1838-1845.
- [24] Danelljan M, Häger G, Khan F S, et al. Accurate scale estimation for robust visual tracking[C]// Proceedings of British Machine Vision Conference. Nottingham, 2014: 65.1-65.11.
- [25] Avidan S. Ensemble tracking[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. San Diego, CA, USA: IEEE, 2005, 29(2): 494-501.
- [26] Dalal N, Triggs B. Histograms of oriented gradients for human detection[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. San Diego, CA, USA: IEEE, 2005.
- [27] Chen D P, Yuan Z J, Hua G, et al. Description-discrimination collaborative tracking[C]// Proceedings of European Conference on Computer Vision. Zurich: IEEE, 2014.
- [28] Felzenszwalb P, Girshick R, McAllester D, et al. Object detection with discriminatively trained partbased models[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(9): 1627-1645.
- [29] Zhang L, Maaten L. Preserving structure in model-free tracking[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014, 36(4): 756-769.
- [30] Ma C, Yang X K, Zhang C Y, et al. Long-term correlation tracking[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Boston, MA USA: IEEE, 2015: 5388-5396.
- [31] Wu Y, Lim J, Yang M H. Online object tracking: a benchmark[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Portland, OR, USA: IEEE, 2013, 9(4): 2411-2418.