

懒交互模式下散乱不规则分块引导的目标跟踪*

刘财兴, 李亚桢, 陈铭钦, 梁云, 甘乙波
(华南农业大学数学与信息学院, 广东 广州 510642)

摘要: 针对目标运动过程中的复杂背景和光照变化等挑战展开研究, 提出基于懒交互和散乱不规则分块的目标跟踪方法, 以提高目标跟踪的准确度和成功率。首先, 采用简单易操作的懒交互方式将目标分成多个散乱分布的不规则分块, 并用核相关滤波对每个分块进行初始化建模; 然后, 基于核相关滤波对每个分块进行跟踪。同时, 为了适应目标和环境的不断变化, 先对分块模型进行简单更新, 当简单更新无法满足目标变换时, 对相关分块进行懒交互式重采样, 以构建更准确的目标分块模型。最后, 根据所有分块在新一帧的位置, 根据霍夫投票确定目标位置。主要创新为: 通过懒交互的方式采样能够保证分块的有效性; 通过对分块进行散乱不规则的采集能够对目标特征进行有效典型的描述。针对 Visual tracker benchmark 的 27 个视频进行测试用于评估不同跟踪方法, 实验结果表明, 本方法处理光照变化、旋转、复杂背景时能够得到更精准的跟踪结果。

关键词: 目标跟踪; 散乱不规则分块; 懒交互重采样; 核相关滤波; 霍夫投票

中图分类号: TP309.7 **文献标志码:** A **文章编号:** 0529-6579 (2018) 04-0062-10

Visual tracking based on lazy interaction and irregular patches

LIU Caixing, LI Yazhen, CHEN Mingqin, LIANG Yun, GAN Yibo

(South China Agriculture University, College of Mathematics and Informatics, Guangzhou 510642, China)

Abstract: A novel visual tracking method based on lazy interaction and irregular patches is proposed, to improving the accuracy and success rate of target tracking, especially on the challenges of background clutters and illumination variation. First, it divides the target into a lot of irregular target patches by using a simple and lazy interactive process, and each target patches is initialized by the kernel correlation filtering. Then, each target patches is tracked based on the kernel correlation filtering. At the same time, the proposed method updates the patches model in a simple way to satisfy the constant changes of the target and the environment. When a simple update fails to deal with the target transformation, the lazy interactive process resamples the related patches and builds a more accurate target patch model. Finally, according to the position of all the target patches on the new frame, the target position is computed by the hough-voting scheme. The main innovation is that sampling the target patches through lazy interaction ensures the effectiveness of describing target patches. And by sampling the irregular patches, it effectively describes the target characteristics. Using the 27 videos of visual tracking benchmark, the quantitative evaluation of different trackers is achieved. Many experiments demonstrate that the proposed method perform is better in dealing with target illumination variation, rotation, background clutters and so on.

Key words: visual tracking; irregular patch; lazy interaction sampling; kernel correlation filtering; hough-voting

* 收稿日期: 2018-01-12

基金项目: 国家自然科学基金 (61772209); 广东省科技计划项目 (2016A050502050)

作者简介: 刘财兴 (1962 年生), 男; 研究方向: 农业信息化技术; E-mail: liu@scau.edu.cn

通信作者: 梁云 (1981 年生), 女; 研究方向: 视频图像处理; E-mail: sdliangyun@163.com

李亚桢 (1992 年生), 男; 研究方向: 视频图像处理; E-mail: liyazhen111@163.com

目标跟踪是计算机视觉领域研究热点之一，通过对视频序列中的目标提取特征，然后根据特征匹配视频流中每帧的位置和目标大小，从而实现目标跟踪。目标跟踪具有广泛的应用前景，如视频监控（实现对目标自动监控）、无人机（实现对目标的自动跟踪拍摄）等。本文采用懒交互将跟踪目标分成多个散乱不规则分块，基于核相关滤波算法实现对各个分块进行建模和跟踪。根据每个分块的置信度值等进行异常判定，采用懒交互方式对异常判定进行重采样。最后运用霍夫投票算法确定跟踪目标在新一帧中的位置。其中，懒交互处理将目标分成多个散乱不规则分块，保证分块对目标特征描述的有效性以及通过对目标典型特征进行多次不规则的采样，能够更好的适应目标状态的变化。与现有跟踪方法相比，本文跟踪算法在旋转、复杂背景、光照变化等挑战的跟踪准确度和成功率具有显著提高。

1 相关工作

目标跟踪是计算机视觉领域研究热点之一。近年来多种目标跟踪方法被提出，如基于相关滤波方法和基于卷积神经网络方法等^[1-11]。但是由于现实情况复杂多变，如目标在跟踪过程中出现遮挡、光照变化、旋转等情况，容易导致跟踪目标丢失。故此，设计出一种鲁棒性很好的跟踪算法是一大挑战。

目标跟踪过程中关键步骤之一是建立符合跟踪目标特征的表现模型。主流的跟踪方法对于跟踪目标的表现模型有两种，一种是用跟踪目标的全局特征描述和表达目标物体^[2,12-15]；一种是将跟踪目标分块多个分块，通过对每个分块进行局部特征来表示跟踪目标^[16-19]。

基于全局特征的目标跟踪，即对跟踪目标建立整体表现模型。在后续跟踪过程中，通过对目标模型进行匹配，实现对目标跟踪效果。如 Zhu 等^[20]从整体特征着手，提出融合梯度特征与颜色特征的跟踪方法，增加目标跟踪的鲁棒性。Bolme 等^[21]提出基于最小平方误差和输出自适应相关滤波器方法，通过对目标整体特征在每一帧上进行匹配，得出目标在每一帧最大值响应值位置，实现目标跟踪，其对目标外观变化例如拉伸收缩有比较强的适应能力。Lukežič 等^[22]通过提取跟踪目标整体的梯度特征和颜色特征提出基于判别相关滤波方法以提高算法的鲁棒性。Wang 等^[23]通过结合支持向量机和相关滤波方法，对跟踪目标进行整体建模，提高

算法对前、背景的区分度，从而提高算法的跟踪准确度。采用整体特征描述跟踪目标，不可避免地会将背景信息融入目标区域中，故此在跟踪过程中容易出现漂移现象，导致跟踪目标丢失。

基于局部特征的目标跟踪，即通过对目标的局部信息进行跟踪来达到对目标整体跟踪的效果。如：Cai 等^[24]提出采用超像素分解目标，通过对计算相邻超像素点像素差值进行连接从而区分背景与目标，增加目标跟踪的准确性。Guo 等^[25]通过学习和训练目标局部特征构建和更新目标表现模型，以解决遮挡或光照等跟踪挑战。Li 等^[16]基于核相关滤波方法提出基于可靠分块的目标跟踪方法，通过将跟踪目标分成 N 个分块，然后对每个分块进行跟踪，来提高算法的鲁棒性。以上方法都是通过规则、随机局部分块对跟踪目标进行描述，在跟踪过程中随着跟踪目标的运动和背景的变化，容易出现分块偏移，导致跟踪目标丢失。

本文在继承基于局部跟踪方法的基础，采用懒交互对目标进行分块建模和更新，进一步保证目标外观的准确性。采用交互方式确定初始化目标，能够很好的确定跟踪目标区域，避免出现冗余区域，但大部分交互方式比较复杂难操作^[26-27]。本文交互方式在目标区域内随意划分分块，简单易操作，故称为“懒交互”。首先在初始化建模阶段，采用懒交互方式将目标分成多个散乱不规则分块。然后运用核相关滤波方法对每个分块进行初始化建模，以此能够有效地控制欠拟合或者过拟合现象。其次对每个分块跟踪过程中进行异常判定，并对异常分块采用懒交互方式进行重采样，从而可以在最大程度上保证分块的有效性。本算法的优点在于：

(i) 采用懒交互方式将目标划分为 N 个散乱不规则分块，简单易操作。

(ii) 对每个分块分别建模，当个别分块出现跟踪异常时，并不会影响整体的跟踪效果，有效地提高了算法的鲁棒性。

(iii) 提出基于懒交互方式目标跟踪方法。在跟踪过程中目标框发生漂移现象时，能及时处理掉异常的块并进行重采，使目标框能够重新正确的对目标进行跟踪，以保证跟踪的延续性，从而保证跟踪算法的准确性。

2 本文系统框架

本文通过懒交互的方式将跟踪目标分成多个散乱不规则分块，提出了一种基于懒交互模式下散乱不规则分块引导的目标跟踪方法。根据实现过程本

文算法主要流程是：首先采用懒交互方式将跟踪目标分成 N 个散乱不规则分块并对每个分块建立初始表观模型；然后根据表观模型和核相关滤波方法对每个分块进行跟踪，并根据跟踪结果确定每个分块在新一帧上的位置；接着对每个分块进行异常判断并对异常分块采用懒交互方式重采样，并对非异常分块更新。最后采用霍夫投票算法^[28]确定跟

踪目标在新一帧上的位置，并更新目标模板进行后续的跟踪。本文算法流程如图 1 所示，包含目标建模阶段、目标跟踪阶段和目标模板更新阶段。

目标建模阶段：在第一帧图像上以标准目标区域为基础，向四周扩展得到一个扩展区域。在扩展区域内采用交互方式随机选取 N 个散乱不规则分块（例如 $N = 25$ ），并提取各个分块中心位置，同

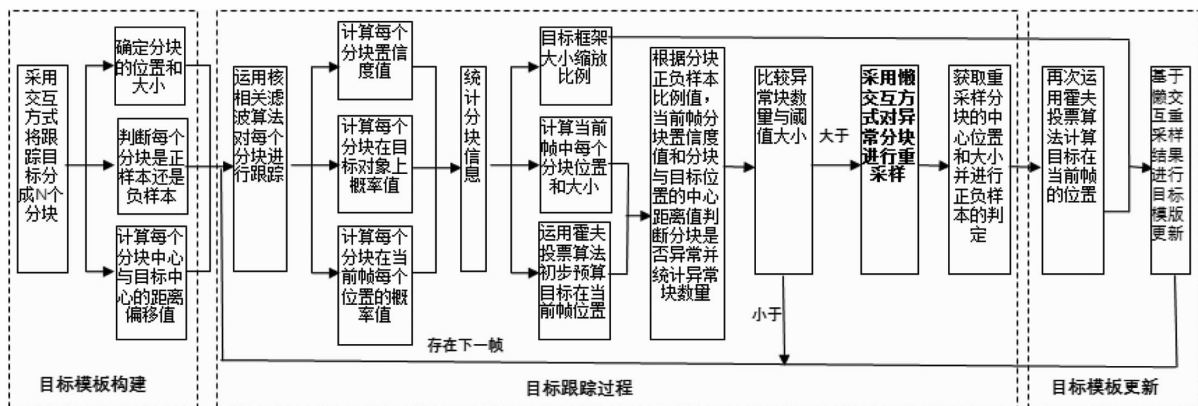


图 1 本文算法总体流程图

Fig. 1 Algorithm flow chart

时计算分块中心与跟踪目标位置的偏移值，以此判定分块正负样本属性，若分块在标准目标区域内则为正样本，否则为负样本。最后，运用核相关滤波方法，构建出每个分块的表观模型。

目标跟踪阶段：输入新一帧图像，对每一个局部分块进行核相关滤波操作，得到每个分块在当前帧出现在每个位置上的概率值和每个分块在目标对象上概率值，从而确定每个分块的作为目标中心位置的的概率值。由此，得到每个分块在当前帧的跟踪位置，再根据所有分块的跟踪位置，并结合霍夫投票算法初步得到跟踪目标位置。其次对每个分块进行异常判定，依据是：分块与跟踪目标中心距离值、分块跟踪中心的置信度值、所有分块的正负样本比例值。然后比较异常块数量与预先设定阈值大小，当异常块数量大于阈值时采用交互方式重采相对应的异常分块数量。最后根据每个分块在新一帧上的中心位置值和每个分块的作为目标中心位置的的概率值运用霍夫投票算法确定目标位置。

目标模板更新：主要为对异常分块重采样的重新建模和对非异常分块的模型更新。重采样分块的建模过程即分块初始化过程。非异常分块的更新为：更新其在新一帧中的参数，建立新的表观模型，再与旧表观模型线性组合。

3 懒交互模式下散乱不规则分块引导的目标跟踪

本章首先介绍如何根据懒交互和散乱不规则分块建立跟踪目标的表观模型，然后结合核相关滤波方法，对每个分块进行跟踪。最后根据所有分块的中心位置，运用霍夫投票计算跟踪目标的跟踪结果，并更新分块的表观模型以精确指导后续帧的跟踪处理。

3.1 建立目标分块的表观模型

建立目标分块表观模型，首先根据懒交互方式选择目标的散乱不规则分块，然后运用核相关滤波建立分块表观模型。

(i) 基于懒交互的目标散乱不规则分块选择。首先，在视频第一帧的目标区域周围，采用懒交互方式选择 N 个散乱布局的分块（如： $N = 25$ ），要求这些分块能够覆盖目标具有典型特征的主要区域，因此分块常是无规则散乱分布且具有不同的尺寸。其次，计算每个分块中心与目标中心的中心偏移值，并据此对每个分块进行正负样本的判定。如图 2 所示，黄色矩形区域为目标区域，橙色区域为基于目标区域的扩展区域，当所取分块的中心位置处于黄色框架内时为正样本，处于橙色框架与黄色

框架之间时为负样本。

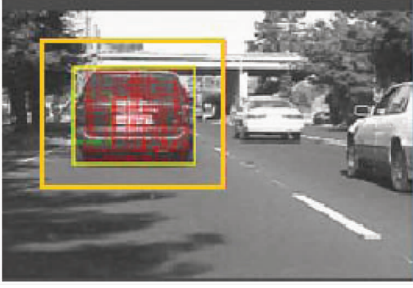


图2 散乱不规则目标分块的采集和划分

Fig.2 Collection and division of scattered and irregular targets

(ii) 基于核相关滤波的分块表观模型。采集目标分块后，运用核相关滤波初始化每个分块。其中，核相关滤波根据分块的表观信息如颜色、梯度等，通过求解岭回归训练目标检测器，它成功地利用了循环矩阵在傅里叶空间可对角化的特性，将矩阵的运算转化为向量的点乘，从而提高了计算效率。本文运用核相关滤波得到每个分块的特征矩阵 X 和用于区分目标与背景的分类器矩阵 α 。

结合上述内容，即可得到每个分块的表观模型。该分块表观模型共包括分块的4部分即：特征矩阵 X ，分类器矩阵 α 、中心偏移、正负样本标记。

3.2 基于核相关滤波的分块跟踪

基于核相关滤波的基本跟踪算子的关键是分类器矩阵的使用和更新。假设第 i 个分块的表观模型的类器矩阵为 α_i ，为跟踪该分块在当前帧提取的目标预测区域的特征表达为 \bar{X}_i ，则该目标预测区域内分块中心的响应值矩阵 $R(\bar{X}_i)$ 如下：

$$R(\bar{X}_i) = \bar{X}_i^T \cdot \alpha_i \quad (1)$$

该矩阵中，响应值越大说明分块中心位于此点的概率越大。核相关滤波最大的特点是跟踪速度快，因此虽然本算法需要对每帧图像中的多个分块进行跟踪，故此能够在最大程度上保证算法的跟踪速度。其中，目标预测区域是根据前一帧目标中心和尺寸得到的。

3.3 基于最优分块的目标跟踪

因本文根据所有散乱不规则分块实现目标跟踪，故需对每个分块进行有效性判定。首先根据分块目标概率值对每个分块进行最优判定（概率值越大，说明离目标中心位置越近）；然后对分块进行异常判定并对异常分块采用懒交互方式进行重采样；最后根据每个分块的位置和其作为目标中心的

概率值，采用霍夫投票算法计算跟踪目标在新一帧的位置。通过将每个分块作为目标中心概率值作为权值，并对异常分块进行及时的重采样操作，保证每个分块的有效性来确保整体跟踪效果的准确性。

3.3.1 最优分块的判定 判定一个分块是否为最优分块的依据是分块在当前帧置信度值和分块在当前帧目标区域上的概率值。结合二者，计算每个分块在当前帧上作为最优分块的概率值。如下式所示：

$$p(z_t | x_t) = p_t(z_t | x_t) p_0(z_t | x_t) \quad (2)$$

其中， z_t 表示当前帧（第 t 帧）， x_t 表示第 t 个分块。所以， $p(z_t | x_t)$ 表示第 t 个分块在当前帧的作为最优分块的概率值， $p_t(z_t | x_t)$ 表示第 t 个分块在当前帧可跟踪的置信度值，它将根据公式（3）计算。 $p_0(z_t | x_t)$ 表示第 t 个分块在当前帧跟踪目标上的概率值，它将根据公式（5）计算。具体而言，最优分块的确定根据如下3步实现：

(i) 计算分块的置信度值。置信度值是基于核相关滤波所得的响应值来计算的。再引入了峰值-旁瓣比来作为衡量分块的置信度值标准。峰值-旁瓣比计算公式：

$$\text{con}(X_i) = \frac{\max(R(X_i)) - \mu_\Phi(R(X_i))}{\sigma_\Phi(R(X_i))} \quad (3)$$

其中， $R(X)$ 是搜索区域上的响应值矩阵， Φ 是在响应峰值的周围区域。 μ_Φ 和 σ_Φ 分别是响应值矩阵中除了 Φ 区域的平均值和标准差。基于 $\text{con}(X_i)$ 的结果，分块的置信度值设置为 $p_t(z_t | x_t) = \text{con}(X_i)^2$ 。

(ii) 计算每个分块在目标区域内的概率值。由于在分块跟踪的过程中，搜索范围包含了背景信息，因此计算过程中需要参考背景的因素，通过对前、背景进行区分可提高整体算法的鲁棒性。本文通过当前分块与分块在上一帧的距离，并借助正样本和负样本的分布信息来计算新块在目标对象上的概率值。计算公式：

$$l(X) = y_i \left(\frac{1}{N^-} \sum_{j \in \Omega^-} V - V^{(j)} \right) - \frac{1}{N^+} \sum_{i \in \Omega^+} V - V^{(i)} \quad (4)$$

其中， $y_i \in \{+1, -1\}$ 表示当前块是正样本还是负样本。 Ω^+ ， Ω^- 分别表示正负样本的集合， N^+ 表示正样本数量， N^- 表示负样本数量。基于 $l(X)$ ，可以计算出这个分块在目标对象上的概率值，计算公式为：

$$p_0(z_t | x_t) = e^{l(X_t)} \quad (5)$$

(iii) 根据公式（2），则可得每个分块作为最

优分块中心的概率值。概率值越大, 离目标中心位置的概率越大。

3.3.2 异常分块的判定及处理 在跟踪过程中随着跟踪目标及背景的变化, 不可避免出现分块跟踪异常的现象, 故需对每个分块进行异常判定并对异常分块进行重采。分块异常判断依据具体如下:

(i) 判断分块是否远离目标。当分块已经超出了目标以及背景的范围, 则需要删除这个分块, 因为这个分块已经不能代表目标或者周围背景的局部区域。

(ii) 判断正负样本比例是否失衡: 在初始分块时是在扩展区域内随机选择分块的, 正负样本的数量也为随机产生。当正样本数量过多时, 容易出现过拟合现象; 当负样本过多时, 则容易跟踪目标丢失。因此, 维持正负样本的比例可保持跟踪算法的稳定性。所以, 当分块的正样本(目标)数或者负样本(背景)数大于某一阈值之后(如: 阈值为 5), 则需要删除掉正样本或者负样本中置信度较小的一些块。

(iii) 所跟踪分块的置信度值是否太低。根据式(1)我们得到每个分块的响应值矩阵, 并运用式(3)和公式 $p_i(z_i | x_i) = \text{con}(X_i)^2$, 得到每个分块的置信度值。当这个分块的置信度值过小, 则说明与原分块的相似度较低, 应该视为异常分块, 应删掉此分块。

通过删除异常分块, 可防止跟踪漂移现象出现; 并通过对异常分块的及时重采样, 提高整体算法的鲁棒性。而未被删除的分块, 在新一帧图像上建立新表观模型并与旧表观模型进行线性组合, 得到最后在新一帧上的表观模型。

3.3.3 异常分块的重采样 由于本文采用懒交互式方式对异常分块进行重采样, 需在保证跟踪准确性的同时最大程度上不影响跟踪速度, 故此对异常分块设定一个重采样判定阈值 m 。当异常分块的数量大于等于 m 时, 则对异常分块进行重采。当阈值设定过高时, 不能对异常分块进行及时的重采样操作, 容易出现跟踪漂移现象; 当阈值设定过低时, 容易导致重采样次数增加, 从而影响跟踪的实时性。在本文设置 $m = 5$, 在保证跟踪准确性的同时, 保证其实时性。其过程如图 3 所示, 其中 (a) 为在进行懒交互式散乱不规则块采集之后目标的跟踪效果; (b) 出现分块漂移现象(左上角红色框); (c) 为进行懒交互式重采样过程; (d) 为进行懒交互式重采样之后目标跟踪效果。

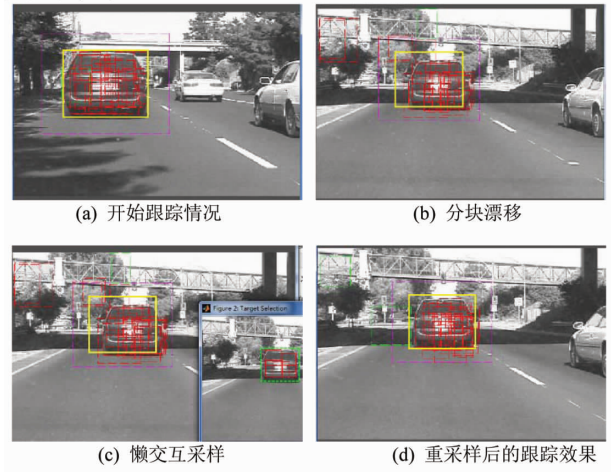


图 3 懒交互式重采样过程(黄色框为目标区域)

Fig. 3 The process of lazy interaction resampling (yellow box for object area)

3.3.4 基于霍夫投票投票的目标跟踪结果计算

在异常分块处理之后, 我们得到了所有的散乱不规则分块, 并依据所以分块采用霍夫投票算法计算跟踪目标在新一帧上的最终位置。将最优分块中心的概率值作为每一个分块的权重, 并与对应的分块位置进行加权求和, 得到跟踪目标在新一帧上的最终位置, 具体计算如式(5)所示。图 4 为霍夫投票示意图。

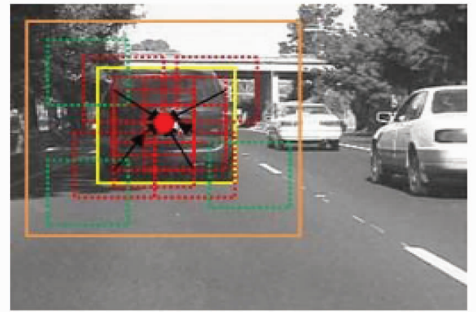


图 4 霍夫投票

Fig. 4 Hough-voting

将最优分块中心的概率值 $p_i(z_i | x_i)$ 作为每一个分块的权重, 并根据式(6)加权计算出跟踪目标中心位置结果。

$$\text{pos}^{\text{tar}} = \sum_{i=1}^N \text{pos}_i \times p_i(z_i | x_i) \quad (6)$$

其中 pos^{tar} 表示跟踪目标在新一帧上的最终位置, pos_i 表示第 i 分块的位置, $p_i(z_i | x_i)$ 表示第 i 个分块作为跟踪目标中心位置的概率值。

3.4 表观模型的更新

表观模型的更新分为 2 种情况：对异常块的更新和对非异常块的更新。

(i) 异常分块表观模型的更新。异常分块都是先删除，然后再对其进行重采样操作。需要对每个新采样的分块利用核相关滤波进行初始化，即对新采样分块建立初始表观模型。

(ii) 非异常分块表观模型的更新。线性组合每个分块的新表观模型和旧表观模型，实现目标分块表观模型的更新。首先，利用核相关滤波和分块的新跟踪结果，计算其新表观模型；然后将该模型与其旧表观模型进行线性组合，具体而言是将描述表观模型的特征矩阵进行组合，最终得到每个分块更新后的表观模型。同时，更新分块与目标中心的偏移值以更新目标的结构信息。

4 实验结果及分析

本文在普通 PC 机（Windows 7 系统，Intel i5-4460CPU，3.20 GHz，4 G 内存）上基于 MATLAB

2012A 实现本文算法。采用 Visual tracking benchmark 平台的 27 组视频序列进行实验，它们包含多种挑战如旋转、快速移动、背景复杂、光照变化等。为验证本算法的优越性，将其与当前主流的 5 种方法（含 KCF^[2]、RPT^[16]、SCM^[29]、Struck^[30]、VTD^[31]）进行对比。结果表明本文方法对平面旋转、外平面旋转、背景复杂、光照变化比较有效。

4.1 跟踪结果定量分析

本文依据中心位置误差精度图和每一帧跟踪区域与目标标准区域的覆盖成功率图对目标跟踪进行定量分析。其中精准度通过跟踪结果区域和理想目标区域的中心距离计算，而成功率则根据两者的重叠率计算^[32]。如图 5 所示，与其它 5 种目标跟踪方法相比，本文方法在中心距离误差精度图和覆盖成功率分别以 0.785 和 0.658 排在第 1。而在光照变化、内平面旋转、复杂背景和外平面旋转（表 1 和表 2）这 4 种情况下，分别以 0.783 和 0.618、0.779 和 0.628、0.759 和 0.624、0.774 和 0.615 排在首位。

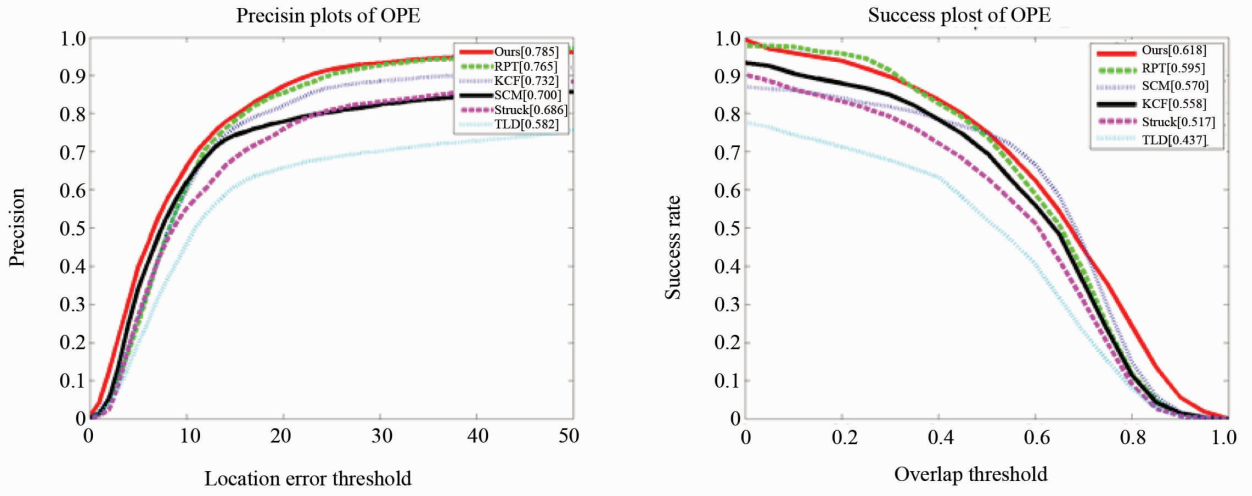


图 5 综合情况

Fig. 5 The total situation

表 1 跟踪精确度（红色为第 1，绿色为第 2、蓝色为第 3）

Table 1 Precision plots of OPE (Red is the first, green is second and blue is third)

Trackers	FM	BC	MB	DEF	IV	IPR	LR	OCC	OPR	OV	SV
OURS	0.711	0.759	0.751	0.735	0.783	0.779	0.759	0.762	0.774	0.693	0.783
KCF	0.684	0.726	0.734	0.706	0.705	0.761	0.726	0.699	0.732	0.569	0.721
RPT	0.698	0.738	0.752	0.731	0.732	0.777	0.738	0.721	0.769	0.451	0.784
Struck	0.618	0.584	0.737	0.601	0.660	0.660	0.584	0.661	0.659	0.493	0.727
TLD	0.587	0.369	0.732	0.438	0.597	0.619	0.369	0.566	0.579	0.674	0.640
SCM	0.549	0.587	0.576	0.569	0.712	0.671	0.587	0.685	0.656	0.480	0.778

表 2 平均成功率
Tabel 2 Success plots of OPE

Trackers	FM	BC	MB	DEF	IV	IPR	LR	OCC	OPR	OV	SV
OURS	0.648	0.624	0.626	0.602	0.618	0.628	0.624	0.562	0.615	0.647	0.593
KCF	0.574	0.590	0.589	0.576	0.531	0.590	0.590	0.514	0.557	0.701	0.508
RPT	0.649	0.592	0.624	0.605	0.572	0.624	0.592	0.566	0.605	0.616	0.591
Struck	0.567	0.480	0.595	0.486	0.500	0.509	0.480	0.490	0.501	0.539	0.504
TLD	0.503	0.282	0.584	0.361	0.446	0.466	0.282	0.444	0.435	0.632	0.475
SCM	0.512	0.488	0.518	0.488	0.601	0.546	0.488	0.564	0.537	0.517	0.633

4.2 目标跟踪结果定性分析

本节主要针对内平面旋转、外平面旋转、背景复杂、光照变化挑战来分析本文方法的有效性，并与现有方法进行对比分析。

4.2.1 内平面旋转 图 5 展示了内平面旋转挑战中两个例子 (football1, dudek) 跟踪效果。由于本文方法通过多个分块表示跟踪目标且在重采样时采用懒交互方式，故当目标在旋转过程中，通过对每个分块进行跟踪以达到对目标的跟踪效果。当出现

异常分块时，通过懒交互方式可以保证重采样分块的有效性，从而确定跟踪的有效性。如在 football1 视频 (第 55 帧, 第 67 帧) 或者 dudek 视频 (第 761 帧) 中，目标在剧烈运动且背景复杂中出现旋转现象，本算法依然能够正确跟踪目标，而其它算法如 RPT 算法，在重采样分块过程中在置信度值最高分块附近进行异常分块重采样，当置信度值最高的分块已发生漂移现象时，则重采样分块也容易出现漂移现象，故容易导致跟踪失败。

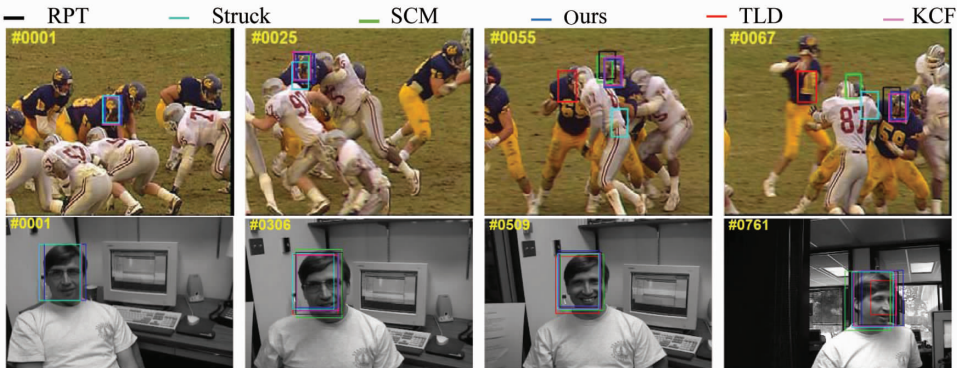


图 6 内平面旋转跟踪结果 (第一行为: football1 序列; 第二行为: dudek 序列)

Fig. 6 The result of in-plane rotation (First line: football; Second line: dudek)

4.2.2 外平面旋转 图 7 展示了外平面旋转挑战中两个例子 (basketball, david3) 跟踪效果。对比分析可知，本文算法通过对每个分块进行建模跟踪，通过每个分块的局部信息来表达整体，当部分分块丢失时可通过其它分块来达到跟踪目标的目的，从而提高整体算法的鲁棒性。在 basketball 视频 (109 帧) 中，SCM、TLD 和 KCF 算法都是通过整体跟踪，故在目标旋转且剧烈运动情况下跟踪失败。在 david3 视频 (54 帧, 118 帧) 中，目标在运动中经过电线杆或者树时被完全遮挡，KCF、Struck 和 SCM 算法采用整体跟踪，故而导致失败。

4.2.3 背景复杂 图 8 展示了背景复杂挑战中两个例子 (carDark, singer2) 跟踪效果。carDark 视

频中，在目标运动路线上背景复杂且出现光照变化情况。第 222 帧左右，目标运动路线出现变化。第 263 帧，如 RPT 算法虽然基于局部分块进行跟踪，但由于分块比较规则，容易融入背景信息，导致出现跟踪漂移现象；而 TLD 算法等背景环境复杂，且采用整体跟踪导致跟踪失败。本文算法在跟踪过程中，出现跟踪分块丢失情况，但通过其余分块调整且通过懒交互方式及时重采样，从而保证后续跟踪的准确性。

4.2.4 光照变化 图 9 展示了光照变化挑战中两个例子 (david, singer1) 跟踪效果。在 david 视频中，目标由光照比较暗的地方走到光照比较亮的地方 (第 387 - 444 帧)，本文算法跟踪效果很稳定。

在 singer1 视频中 (第 195 帧、第 286 帧), 光照变化很频繁, 本文通过对散乱不规则分块跟踪性且通

过懒交互重采样方式, 最大程度保证分块有效性, 在后续模版更新及跟踪时, 可维持算法的鲁棒性。

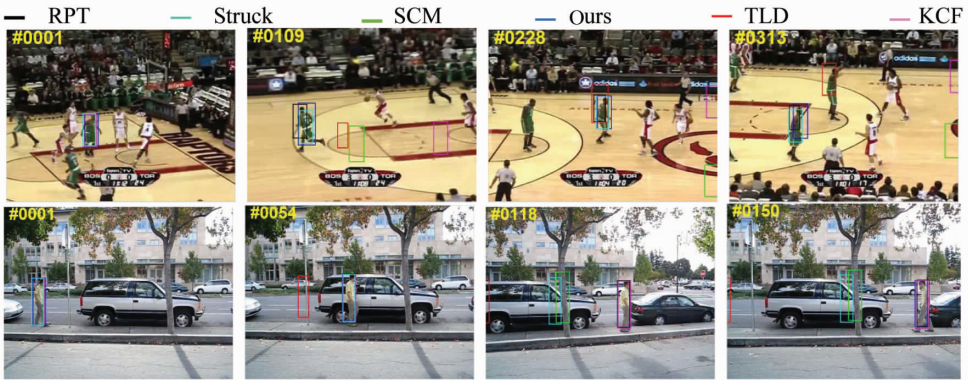


图 7 外平面旋转跟踪结果 (第一行为: basketball 序列; 第二行为 david3 序列)
Fig. 7 The result of out-of-plane rotation (First line: basketball; Second line: david3)



图 8 背景复杂跟踪结果第一行为: carDark 序列; 第二行为 singer2 序列)
Fig. 8 The result of background clutters (First line: carDark; Second line: singer2)

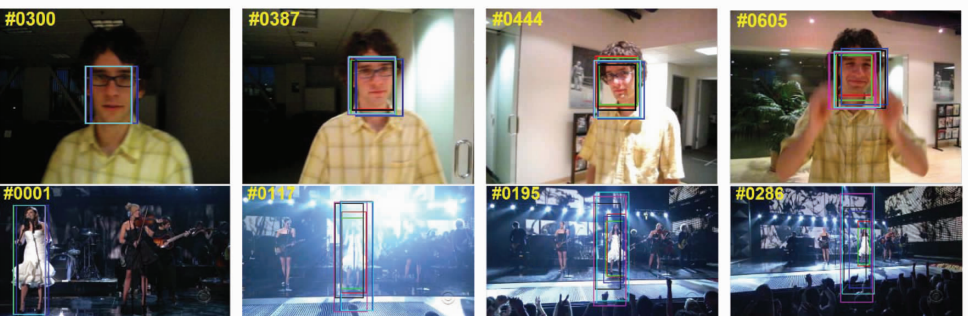


图 9 光照变化跟踪结果 (第一行为: david 序列; 第二行为 singer1 序列)
Fig. 9 The result of illumination variation (First line: david; Second line: singer1)

4.3 实时性分析

如表 3 所示, 列举了本文算法与其它四种算法的平均速度对比数据。由表中可知, 本文算法的运行速度为 7.8 Fps, 远小于 KCF 算法 (195.1 Fps), 但大于同样通过对目标分块跟踪以达到对目标整体跟踪的 RPT 算法 (3.59 Fps)。

表 3 运行平均速度对比

Tabel 3 Running average speed contrast

Trackers	Ours	KCF	RPT	Struck	SCM
mFps	7.80	195.10	3.59	5.04	0.32

5 结 语

本文针对在视频跟踪过程中出现背景复杂、光照变化等挑战展开研究,提出了一种懒交互模式下散乱不规则分块引导的目标跟踪方法。该方法首先将跟踪目标区域通过懒交互方式划分为 N 个散乱不规则分块,然后分别对 N 个分块进行初始化建模,并引入核相关滤波算法提高跟踪效率。针对异

常分块,本文提出一种懒交互方式方法对异常分块进行重采。最后采用霍夫投票算法确定跟踪目标在新一帧上的最终位置。通过在 benchmark 上对大量视频序列进行跟踪的结果表明:与现有主流跟踪算法相比,本算法对多种挑战(如背景复杂、光照变化、旋转等)情况下,本方法的跟踪结果具有更高的精度。

参考文献:

- [1] DANELLJAN M, SHAHBAZ KHAN F, FELSBERG M, et al. Adaptive color attributes for real-time visual tracking [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014: 1090 – 1097.
- [2] 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.
- [3] DANELLJAN M, HÄGER G, KHAN F, et al. Accurate scale estimation for robust visual tracking [C]// British Machine Vision Conference, Nottingham, September 1 – 5, 2014. BMVA Press, 2014.
- [4] BERTINETTO L, VALMADRE J, GOLODETZ S, et al. Staple: Complementary learners for real-time tracking [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 1401 – 1409.
- [5] ZHANG K, ZHANG L, LIU Q, et al. Fast visual tracking via dense spatio-temporal context learning [C]// European Conference on Computer Vision, Springer, Cham, 2014: 127 – 141.
- [6] BOLME D S, BEVERIDGE J R, DRAPER B A, et al. Visual object tracking using adaptive correlation filters [C]// Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on. IEEE, 2010: 2544 – 2550.
- [7] WANG L, OUYANG W, WANG X, et al. Visual tracking with fully convolutional networks [C]// Proceedings of the IEEE International Conference on Computer Vision, 2015: 3119 – 3127.
- [8] MA C, HUANG J B, YANG X, et al. Hierarchical convolutional features for visual tracking [C]// Proceedings of the IEEE International Conference on Computer Vision, 2015: 3074 – 3082.
- [9] WANG L, OUYANG W, WANG X, et al. Stct: Sequentially training convolutional networks for visual tracking [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 1373 – 1381.
- [10] QI Y, ZHANG S, QIN L, et al. Hedged deep tracking [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 4303 – 4311.
- [11] HONG S, YOU T, KWAK S, et al. Online tracking by learning discriminative saliency map with convolutional neural network [C]// International Conference on Machine Learning, 2015: 597 – 606.
- [12] WANG D, LU H, YANG M H. Least soft-threshold squares tracking [C]// IEEE Conference on Computer Vision and Pattern Recognition, 2013: 2371 – 2378.
- [13] 王美华,梁云,刘福明,等. 部件级表观模型的目标跟踪方法[J]. 软件学报, 2015,26(10): 2733 – 2747.
WANG M H, LIANG Y, LIU F M, et al. Object tracking method of component level apparent model [J]. Journal of Software, 2015, 26(10): 2733 – 2747.
- [14] 夏侯凯顺,严娟,叶小朋,等. 基于 Kalman 滤波的无线传感器网络多目标跟踪[J]. 中山大学学报(自然科学版), 2014, 53(2): 18 – 22.
XIAHOU K S, YAN J, YE X P, et al. Multi object tracking in wireless sensor networks based on Kalman filtering [J]. Acta Scientiarum Naturalium Universitatis Sunyatseni, 2014, 53(2): 18 – 22.
- [15] ZHU G, PORIKLI F, LI H. Beyond local search: Tracking objects everywhere with instance-specific proposals [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 943 – 951.
- [16] LI Y, ZHU J, HOI S C H. Reliable patch trackers: Robust visual tracking by exploiting reliable patches [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 353 – 361.
- [17] HE S, YANG Q X, LAU R W H, et al. Visual tracking via locality sensitive histograms [C]// Computer Vision and Pattern Recognition, IEEE, 2013: 2427 – 2434.
- [18] JIA X, LU H, YANG M H. Visual tracking via adaptive structural local sparse appearance model [C]// Computer Vision and Pattern Recognition. IEEE, 2012: 1822 –

- 1829.
- [19] KWON J, LEE K M. Highly nonrigid object tracking via patch-based dynamic appearance modeling [J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2013, 35(10):2427–2441.
- [20] ZHU G, WANG J, WU Y, et al. MC-HOG correlation tracking with saliency proposal [C]// *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence*. AAAI Press, 2016: 3690–3696.
- [21] BOLME D S, BEVERIDGE J R, DRAPER B A, et al. Visual object tracking using adaptive correlation filters [C]// *IEEE Conference on Computer Vision and Pattern Recognition*, 2010: 2544–2550.
- [22] LUKEŽIČ A, VOJÍŘ T, ČEHOVIN L, et al. Discriminative correlation filter with channel and spatial reliability [C]// *IEEE Conference on Computer Vision and Pattern Recognition*, 2017:4847–4856.
- [23] WANG M, LIU Y, HUANG Z. Large margin object tracking with circulant feature maps [J]. *arXiv preprint*, arXiv:1703.05020, 2017.
- [24] CAI Z, WEN L, YANG J, et al. Structured visual tracking with dynamic graph [C]// *Asian Conference on Computer Vision*. Berlin Heidelberg: Springer, 2012: 86–97.
- [25] LI A, TANG F, GUO Y, et al. Discriminative non-orthogonal binary subspace tracking [C]// *European Conference on Computer Vision Conference on Computer Vision*. Springer-Verlag, 2010: 258–271.
- [26] WANG Y, LUO Z, JODOIN P M. Interactive deep learning method for segmenting moving objects [J]. *Pattern Recognition Letters*, 2017, 96: 66–75.
- [27] MEKA A, FOX G, ZOLLHÖFER M, et al. Live user-guided intrinsic video for static scenes [J]. *IEEE Transactions on Visualization and Computer Graphics*, 2017, 23(11): 2447–2454.
- [28] GODEC M, ROTH P M, BISCHOF H. Hough-based tracking of non-rigid objects [J]. *Computer Vision and Image Understanding*, 2013, 117(10): 1245–1256.
- [29] ZHONG W, LU H, YANG M H. Robust object tracking via sparsity-based collaborative model [C]// *IEEE Conference on Computer Vision and Pattern Recognition*, 2012: 1838–1845.
- [30] HARE S, SAFFARI A, TORR P H S. Struck: Structured output tracking with kernels [C]// *IEEE International Conference on Computer Vision*, 2012:263–270.
- [31] KALAL Z, MATAS J, MIKOLAJCZYK K. Pn learning: Bootstrapping binary classifiers by structural constraints [C]// *IEEE Conference on Computer Vision and Pattern Recognition*, 2010: 49–56.
- [32] WU Y, LIM J, YANG M H. Online object tracking: A benchmark [C]// *IEEE Conference on Computer Vision and Pattern Recognition*, 2013: 2411–2418.