

一种健壮的超像素跟踪算法^①

郭利^{1,2}, 周盛宗^{1,2,3}, 付璐斯^{1,3}, 于志刚^{1,3}

¹(中国科学院 福建物构所, 福州 350002)

²(福建师范大学 数学与计算机科学学院, 福州 350117)

³(中北大学 计算机与控制工程学院, 太原 030051)

摘要: 在目标跟踪中, 传统的超像素跟踪算法在发生遮挡等情况后, 会将非目标超像素标记为目标加入到特征空间. 在对候选样本置信度计算中, 利用特征空间中最近邻超像素来划定样本中超像素的簇归属会产生错误; 而依据的近邻超像素数量过多时, 又会造成分类误差的积累. 为解决上述问题, 本文提出一种健壮的超像素跟踪算法. 本算法以贝叶斯算法为框架, 首先, 将前几帧进行超像素切割, 提取特征并使用均值漂移聚类算法和基于超像素的外观表示模型进行分类和计算类置信度, 放入特征空间中. 其次, 根据接下来几帧的平均中心误差确定最佳近邻数目. 最后, 在跟踪过程中, 对获取帧的指定区域进行超像素切割, 提取特征、进行软分类和计算置信度; 根据上一帧目标位置进行高斯采样, 累加样本内超像素置信度, 获得样本置信度; 在发生严重遮挡时, 不进行滑动窗口更新和外观模型修改, 使用当前模型继续跟踪. 与传统的最近邻超像素算法相比, 本算法能够有效提升跟踪成功率和降低平均中心误差.

关键词: 最佳近邻数目; 软分类; 目标跟踪; 均值漂移算法; 置信度; 遮挡

引用格式: 郭利, 周盛宗, 付璐斯, 于志刚. 一种健壮的超像素跟踪算法. 计算机系统应用, 2017, 26(12): 130-136. <http://www.c-s-a.org.cn/1003-3254/6120.html>

Robust Superpixels Tracking Method

GUO Li^{1,2}, ZHOU Sheng-Zong^{1,2,3}, FU Lu-Si^{1,3}, YU Zhi-Gang^{1,3}

¹(Fujian Institute of Research on the Structure of Matter, Chinese Academy of Sciences, Fuzhou 350002, China)

²(College of Mathematics and Computer Science, Fujian Normal University, Fuzhou 350117, China)

³(School of Computer and Control Engineering, North University of China, Taiyuan 030051, China)

Abstract: During the object tracking, when occlusion occurs, the traditional superpixel tracking algorithm will add the superpixels of the non-target area into the feature space. In the calculation of the candidate sample confidence, the nearest neighbor superpixel in the feature space is used to delimit the cluster attribution of the superpixels in the sample, and the accumulation of the classification error is caused by the excessive number of neighboring superpixels. To solve the problems above, we propose a robust superpixels tracking method. This algorithm uses Bayesian algorithm as the framework. Firstly, we slice the first few frames into superpixels, extract the feature, use the mean shift clustering algorithm and representation model based on superpixel to classify and calculate the class confidence value, and put the feature into feature space. Secondly, the suitable numbers of neighbors can be found with the mean center error of next few frames. Last but not least, during the tracking process, the superpixel is segmented in the specified area of the acquired frame, to extract the feature. The cluster is confirmed with soft classification and the confidence value is calculated. According to the previous frame target position, the Gaussian sampling is collected. We can obtain the sample confidence value with the accumulation of the confidence value. In case of severe occlusion, the sliding window update

① 收稿时间: 2017-03-16; 修改时间: 2017-04-05; 采用时间: 2017-04-17

and the appearance model modification are not carried out, and we continue to use the current model to track. Compared with the traditional tracking algorithm based on nearest superpixel, the algorithm can effectively improve the tracking success rate and reduce the average center errors.

Key words: best number of nearest neighbors; soft classification; object tracking; meanshift; confidence; occlusion

运动目标跟踪是机器视觉研究的一个基本课题,在视频监控、运动分析、人机交互和自动驾驶等领域中发挥着重要作用.目标跟踪是从一系列图像中,确定目标在每帧图像中的位置、方向、面积、形状等.由于受光照强弱、相机颠簸、物体遮挡、形态变化、速度调整等因素的影响,目标外观发生了很大的变化,从而使得目标跟踪非常有难度. A.Yilmaz 等^[1]认为目标跟踪主要分为点跟踪、核跟踪、剪影跟踪三类.点跟踪主要包括确定法和统计法,代表性的方法有 MGE^[2]、GOA^[3]、卡尔曼滤波^[4]、JPDAF^[5]、PMHT^[6]等;核跟踪主要包括基于模板和密度的外观模型、多角度外观模型等,前者的代表性方法有 Mean-Shift^[7]、KLT^[8]、分层^[9],后者的代表性方法有特征值跟踪^[10]、支持向量机跟踪^[11];剪影跟踪主要包括轮廓演变、形状匹配,前者的代表性方法有状态空间模型^[12]、变量方法^[13]、启发方法^[14].这些数算法大都是围绕着目标表示、特征选择、运动建模、外观表示、形态变化等五个方面进行改进,在特定场景下取得了不错的效果.在目标运动中,自适应外观算法^[15-20]能够很好的表示外观的变化,但基于低层特征的自适应外观算法在前后帧的关联上表现很差;Kwon 等^[21]提出的基于多运动模型和改进观测模型的粒子滤波框架能较好的处理遮挡、尺度、光照和尺度等因素引起的外观变化,由于未区分目标、背景,会出现跟踪目标丢失. Ren 等^[22]提出的超像素跟踪算法,使用 Delaunay 三角化和条件随机场单独对每帧完整图做区域匹配,能够对前景、背景有限分割,但仍然无法处理跟踪过程中的遮挡、复杂背景等问题. Yant 等^[23]提出的健壮超像素跟踪算法,在有限遮挡、光照变化、姿态变化、复杂背景等情况下,都取得不错效果,但由于遮挡等情况下,划定到特征空间中的超像素在目标、背景划分及簇归属上会存在一些错误,仅根据特征空间中离其最近的超像素的簇归属来划定簇是错误的;同样,依据的超像素数量过多时,又会造成分类误差的积累,有限数目超像素来确定当前超像素簇归属会更有效.因此,提出一种健壮的超像素跟踪算法,该算法根据历史样本匹配度寻找最佳近邻数目,使用贝叶斯框架构造联合外观模型,使用基于超像素的判别型外观模型将前景对象从背景中区分开来.

1 算法的主要部分

1.1 确定最佳近邻个数

与传统的基于近邻得票数多少来判定类型的 K-近邻判别法不同的是,本文使用基于历史样本的自适应近邻判别算法^[24].已知基空间中有 n 个样本 m 个类, θ_j 表示第 j 个样本 X_j 所属的类别, $X(=x)$ 与基空间中的样本 X_1, X_2, \dots, X_n 的欧式距离,从小到大排列为: $D(x, X_{n1}) \leq D(x, X_{n2}) \leq \dots \leq D(x, X_{nm})$, 对应的类别为 $\theta_{n1}, \dots, \theta_{nk}, \dots, \theta_{nn}$, 把 $1, \dots, m$ 在 $\theta_{n1}, \dots, \theta_{nk}$ 中得票最多的记为判别值 $\theta_n^{(k)}$.

如图 1, $K=5$ 时,黄色椭圆的最近 5 个邻居是 2 个绿色正方形和 3 个红色三角形, K-近邻判别法判定黄色椭圆与红色三角形同类; $K=9$ 时,黄色椭圆的最近 9 个邻居是 5 个绿色正方形、3 个红色三角形和 1 个深红梯形, K-近邻判别法判定黄色椭圆与绿色正方形同类.

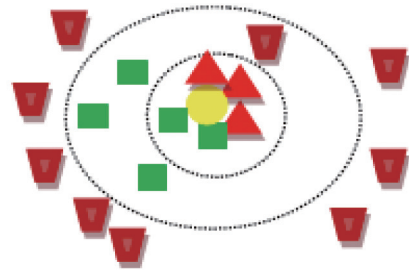


图 1 不同近邻对分类的影响

同理, $\theta_{jn}^{(k)}$ 表示利用除第 j 个样本外的历史样本的 k -近邻判别值(得票率最高) 来确定第 j 个样本的类; 通过式 (1) 获得最大分类准确率时近邻数目 k^* .

$$k^* = \arg \max_{k'} \{E_n^{(k')}\} = \max_{1 \leq k \leq (\log n)^{\sigma_0}} E_n^{(k)}, \quad 0 < \sigma_0 < 1 \quad (1)$$

其中, $E_n^{(k)} = \sum_{j=1}^n I(\theta_{jn}^{(k)} = \theta_j)$.

1.2 贝叶斯跟踪框架

给定 t 时刻的状态 X_t 、所有的观测值 $Y_{1:t}$, 设定归一化项为 α , 计算最大后验分布^[23], 具体如式 (2).

$$p(X_t | Y_{1:t}) = \alpha p(Y_t | X_t) \int p(X_t | X_{t-1}) p(X_{t-1} | Y_{1:t-1}) dX_{t-1}, \quad (2)$$

其中,目标状态 $X_t = (X_t^c, X_t^{sx}, X_t^{sy})$, X_t^c 表示目标的中心位置, X_t^{sx}, X_t^{sy} 分别表示 x 方向、 y 方向的尺度, 观测模型是 $p(Y_t|X_t)$, 运动模型是 $p(X_t|X_{t-1})$. 在给定观测值的情况下, 一个健壮的可区分模型可以用来计算属于目标或背景的可能性. 特定目标的候选 X_t 的观测值评估 $p(Y_t|X_t)$ 与置信度 $\widehat{C}(X_t)$ 成比例^[23]:

$$p(Y_t|X_t) \propto \widehat{C}(X_t), \quad (3)$$

设定 $X_t^{(i)}$ 为状态 X_t 时的第 i 个样本, 则:

$$\widehat{X}_t = \arg \max_{X_t^{(i)}} p(X_t^{(i)}|Y_{1:t}), \quad \forall i = 1, \dots, N. \quad (4)$$

1.3 基于超像素的外观区分模型

首先定义属于目标的像素为 1, 属于背景的像素为-1; 然后对每一帧进行超像素切割, 并进行聚类; 根据超像素中的像素计算超像素的置信度, 根据簇中的超像素计算簇的置信度, 根据帧中超像素所属的簇计算出该帧的置信度.

$S^+(i)$ 表示在状态覆盖目标区域的簇面积, $S^-(i)$ 表示状态 X_t 时覆盖目标区域外的簇面积, 有目标置信度^[23]:

$$C_i^c = S^+(i) - S^-(i) / S^+(i) + S^-(i), \quad \forall i = 1, \dots, n. \quad (5)$$

正值越大, 簇属于目标的置信度越高, 反之亦然.

1.4 基于 k 近邻的置信度计算

当新帧到来时, 使用高斯采样提取目标的周边区域, 根据当前超像素与簇空间中超像素的距离来计算置信度, 并确定置信度最高的样本为跟踪目标. 在遮挡等情况下, 超像素在目标、背景的划分上较难, 仅依据空间中最近邻的超像素来划定簇归类不准确. 因为置信度与两个因素有关: 距离较近的超像素所属的簇、在特征空间中超像素与簇中心的距离. 取定近邻超像素的个数为 $k=k^*$, 距离分别为 D_1, D_2, \dots, D_k , 则:

$$\begin{aligned} M_1 : M_2 : \dots : M_k &= D_1 : D_2 : \dots : D_k; \quad \sum_{i=1}^k 1/M_i = 1; \\ \sum_{i=1}^k 1/M_i &= (1/M_1) \times [1 + D_1(1/D_2 + 1/D_3 + \dots + 1/D_k)] = 1; \\ M_1 &= 1 + D_1(1/D_2 + 1/D_3 + \dots + 1/D_k); \\ M_i &= D_1 \times M_1 / D_i = D_1 \times [1 + D_1(1/D_2 + 1/D_3 + \dots + 1/D_k)] / D_i; \\ w(r, i) &= \exp(-\lambda_d / r_c(i) \times \|f_r^t - f_c(i)\|_2), \quad (6) \\ \forall r &= 1, \dots, N_t, \quad i = 1, \dots, N_t, \end{aligned}$$

$$C_r^s = \sum_{i=1}^k (1/M_i) \times w(r, i) \times C_i^c, \quad \forall i = 1, \dots, N_t. \quad (7)$$

其中, $M_i, i = 1, \dots, k$ 表示空间中第 i 个超像素的贡献度,

f_r^t 是 $\text{sp}(t, r)$ 的特征, 第 t 帧的第 r 个超像素; $f_c(i)$ 是 $\text{sp}(t, r)$ 所属的簇中心; $r_c(i)$ 是 $\text{clst}(i)$ 簇半径; λ_d 是归一化因子, 默认取 2; w 表示基于 f_r^t 和 $f_c(i)$ 间距离的权重; C_r^s 表示第 t 帧第 r 个超像素基于超像素空间中距离较近的 k 个超像素的贡献度的置信度.

1.5 观察和运动模型

设定运动模型为高斯分布^[23],

$$p(X_t|X_{t-1}) = N(X_t; X_{t-1}, \psi), \quad (8)$$

ψ 为对角协方差矩阵, 元素主要是位置和尺度的标准偏移. 规范化所有候选图片区域为与 X_{t-1} 相关的图集 $\{M_k\}_{k=1}^N$. 设定 $v_k(i, j)$ 为 $X_t^{(k)}$ 的规范化置信图 M_k 在位置 (i, j) 处的值, 则在状态 $X_t^{(k)}$ 时的置信度^[23]为:

$$C_k = \sum_{(i,j) \in M_k} v_k(i, j). \quad (9)$$

为了将目标尺度变化^[23]考虑进去, 设定 $S(X_{t-1})$ 代表目标状态 X_{t-1} 时的面积大小, $S(X_t^{(k)})$ 代表目标状态 $X_t^{(k)}$ 时的面积大小, 则:

$$\widehat{C}_k = C_k \times [S(X_t^{(k)}) / S(X_{t-1})], \quad \forall k = 1, \dots, N, \quad (10)$$

观测模型^[23]:

$$p(Y_t|X_t^{(k)}) = \widehat{C}_k, \quad \forall k = 1, \dots, N, \quad (11)$$

1.6 针对遮挡和漂移的在线更新

使用滑动窗口机制进行跟踪, 每 U 帧, 添加一个新帧到 H 帧序列中, 并删除序列中的第 k 帧 ($k < H$).

使用 O_t 计算遮挡指标, 设定时刻 t 的置信度 $C_k \in [-S(X_t^{(k)}), S(X_t^{(k)})]$, μ_t 为保存的 H 帧序列中目标评估的规范化置信度的平均值, 当前帧的目标评估 \widehat{X}_t 为最后一帧的目标评估值 X_{t-1} . 遮挡指标^[23]为:

$$O_t = \mu_t - \max(\{C_k\}_{k=1}^N) / S(X_{t-1}), \quad (12)$$

大于零时, 表示当前帧的平均准确率评估的置信度比保存序列中的规范化置信度的平均值小, 超过一定阈值时, 候选区域可能属于背景区域, 或发生严重遮挡.

2 算法的实现过程

2.1 初始化

for $t = 1$ to m (设定 $m=4$)

1) 初始化 t 帧的各个参数.

2) 将 X_t 周边区域切割为 N_t 个超像素, 提取特征 $\{f_r^t\}_{r=1}^{N_t}$ 用于训练.

End

获取特征池 $F = \{f_r^t | t = 1, 1, \dots, m; r = 1, \dots, N_t\}$, 用均值漂移算法聚类, 使用公式 (5) 获取基于超像素的可区

分外观模型。

2.2 学习近邻数目

for $t = m+1$ to n (设定 $n=15$)

1) 切割 X_t 周边区域为 N_t 个超像素, 提取特征. 使用公式 (6)、(7) 计算目标背景置信度.

2) 用置信图获取 N 个候选状态 $\{X_t^{(i)}\}_{i=1}^N$.

3) 使用公式 (8) 计算运动参数 $p\{X_t^{(i)}|X_{t-1}\}$, 公式 (9)、(10)、(11) 计算置信度 $p\{Y|X_t^{(i)}\}$.

4) 依据公式 (1), 获取最佳后验时的近邻数目 k .

5) 使用公式 (12) 来探测严重或完全遮挡.

6) 每隔 U 帧, 添加一帧到更新序列.

7) 每隔 W 帧, 更新一次外观模型.

2.3 跟踪

for $t = n+1$ to 帧数

1) 切割 X_t 周边区域为 N_t 个超像素, 提取特征. 使用公式 (6)、(7) 计算值目标背景置信度.

2) 用置信图获取 N 个候选状态 $\{X_t^{(i)}\}_{i=1}^N$.

3) 使用公式 (8) 计算运动参数 $p\{X_t^{(i)}|X_{t-1}\}$, 公式 (9)、(10)、(11) 计算相似性 $p\{Y|X_t^{(i)}\}$.

4) 获取最大后验时的状态 \hat{X}_t .

5) 使用公式 (12) 来探测严重或完全遮挡.

6) 每隔 U 帧, 添加一帧到更新序列.

7) 每隔 W 帧, 更新一次外观模型.

End

3 实验结果及分析

实验在 CPU 为 3.6 GHZ、内存为 8 G 的 PC 机上进行, 采用 Matlab2012b、Microsoft Visual Studio 2010 混合编程实现.

3.1 实验设置

HIS 颜色空间降低了像素受光照变化的影响, 在区分不同超像素方面, 比其他颜色空间的区分能力更强, 使用 HIS 颜色空间中的规范化图表作为每个超像素的特征. 使用 SLIC 算法切割图片时, 空间近似权值设置为 10, 超像素数目设置为 300, 均值漂移聚类算法的尺度设置为 0.18.

在初始阶段, 为了获取训练集, 前 4 帧的目标区域可以通过目标定位器或手动获取. 按照经验, 设置 $H=15$, $U=1$, $W=10$, 速度变化 $\sigma_c=7.6$, 尺度变化 $\sigma_s=7$. 为简化遮挡探测, 设置 $\mu_c=0.5$.

3.2 实验结果及分析

表 1 对比了 IVT^[25](不断学习的健壮跟踪)、Frag^[18](使用积分直方图的基于碎片的健壮跟踪算法)、

MIL(在线的多实例学习的视觉跟踪)、VTD^[26](视觉跟踪分解)、L1^[27](使用加速接近梯度算法的实时健壮 L1 跟踪)、TLD^[28](基于机构约束的增强二进制分类器)、Struck^[29](基于核的机构性输出跟踪)、SPT^[23](基于超像素的目标跟踪)、KSPT(基于有限学习的软分类超像素跟踪算法) 在 Lemming、liquor、basketball、bolt 等数据集上中心定位错误的帧数上, KSPT 都是最少的.

表 1 中心定位错误的帧数

序列	IVT	Frag	MIL	VTD	L1	TLD	Struck	SPT	KSPT
lemming	14	84	14	98	182	104	134	7	5
liquor	238	31	165	155	80	28	124	9	8
basketball	120	14	104	11	100	170	153	6	5
bolt	382	100	380	14	369	90	387	6	6

表 2 对比了 IVT^[25]、Frag^[18]、MIL、VTD^[26]、L1^[27]、TLD^[28]、Struck^[29]、SPT^[23]、KSPT 在 Lemming、liquor、basketball、bolt 等数据集上成功跟踪到的帧数, KSPT 都是最多的.

表 2 成功跟踪到的帧数

序列	IVT	Frag	MIL	VTD	L1	TLD	Struck	SPT	KSPT
lemming	1053	680	1105	470	226	361	652	1277	1290
liquor	400	1377	353	471	988	1398	405	1689	1715
basketball	80	512	204	619	30	46	85	695	701
bolt	5	33	12	199	14	49	9	231	240

表 3 对比了 SPT^[23]、KSPT 在 surfing1、singer1、woman、girl_move 等数据集上的成功率, KSPT 比 SPT 都有了很大提升. 这是因为超像素空间中, 各类在簇的划分上, 并不绝对准确, 最近邻超像素跟踪算法只依据与超像素空间最近邻的超像素进行分类, 易发生分类错误. 而通过训练过程, 学习出最佳的近邻数目, 引导超像素分类, 一定程度增加了待分类的超像素的容错程度, 提升了超像素置信度计算的准确度, 从而提高算法对目标的跟踪能力.

表 3 成功率 (%)

序列	SPT	KSPT
Surfing1	58.6792	81.2576
Singer1	76.7320	86.2078
woman	49.9538	61.2578
girl_move	61.3780	83.7532

图 2 中, 不同近邻个超像素的 SPT 在 Lemming 数据集第 16~25 帧上的位置误差情况, 1、10、200 近邻个超像素的 SPT 对应位置总误差 59.3936、36.3729、70.3006 像素; 10 近邻个超像素 (*型实线) 的 SPT 表现最好, 其他近邻的 SPT 总误差相对较大. 在此 10 帧的

跟踪中, 10 近邻超像素的 SPT 表现更为稳定, 误差相对最小. 根据该 10 帧, 推测出 KSPT 中最佳 K=10; 而事实上, 在 26~800 帧上, 10 近邻的 SPT 表现也最好, 其中 370~430、460~600 帧表现最明显, 见图 3、图 4.

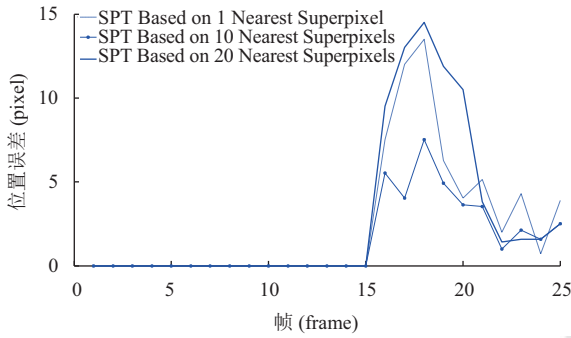


图 2 不同近邻个超像素的 SPT 的位置误差

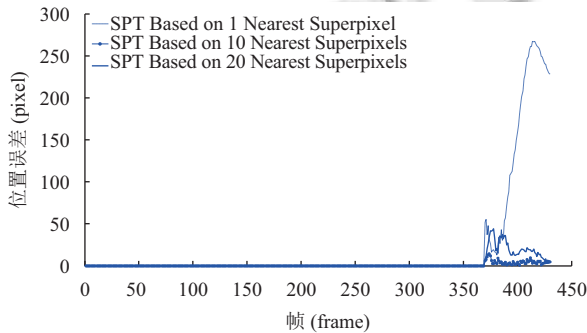


图 3 不同近邻个超像素的 SPT 的位置误差

图 3 中, 不同近邻个超像素的 SPT 在 Lemming 数据集第 370~430 帧上位置误差, 1、10、200 近邻个超像素的 SPT 对应位置总误差 9088.5、275.3、12043 像素; 10 近邻个超像素 (*型实线) 的 SPT 表现最好, 其他近邻的 SPT 总误差相对较大.

图 4 中, 不同近邻个超像素的 SPT 在 Lemming 数据集第 460~600 帧上的位置误差情况, 1、10、200 近邻个超像素的 SPT 对应位置误差 31904、779、2594 像素.

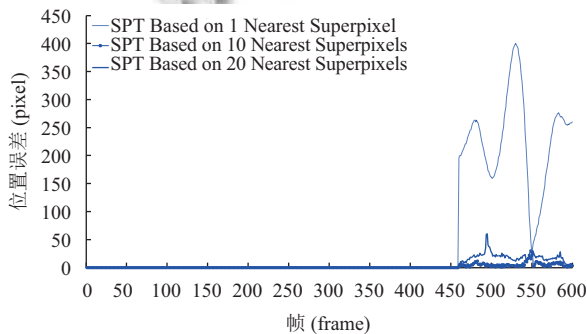


图 4 不同近邻个超像素的 SPT 的位置误差

图 5 中, 不同近邻个超像素的 SPT 在 skating1 数据集第 16~25 帧上位置误差, 1、10、20 近邻个超像素的 SPT 对应位置总误差为 60.8563、40.4681、56.9368 像素; 10 近邻个超像素 (*型实线) 的 SPT 表现最好, 其他近邻的 SPT 总误差相对较大. 在此 10 帧的跟踪中, 10 近邻超像素的 SPT 表现更为稳定, 误差相对最小. 根据该 10 帧, 推测出 KSPT 中最佳 K=10; 而事实上, 10 近邻超像素的 SPT 在 26~300 帧上, 10 近邻的 SPT 表现也很好, 见图 6、图 7. 在 75~230 帧上, 20 近邻的 SPT 发生跟踪漂移, 跟踪效果很差, 见图 6; 260~300 帧, 1 近邻的 SPT 发生跟踪漂移, 跟踪效果也很差, 见图 7.

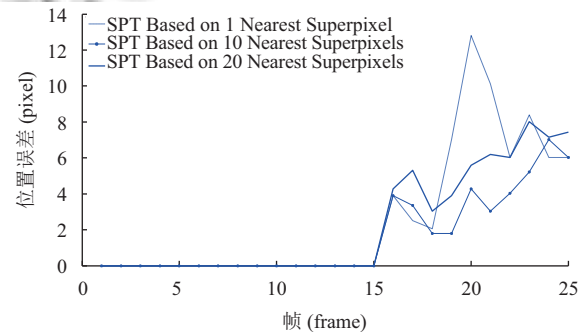


图 5 不同近邻个超像素的 SPT 的位置误差

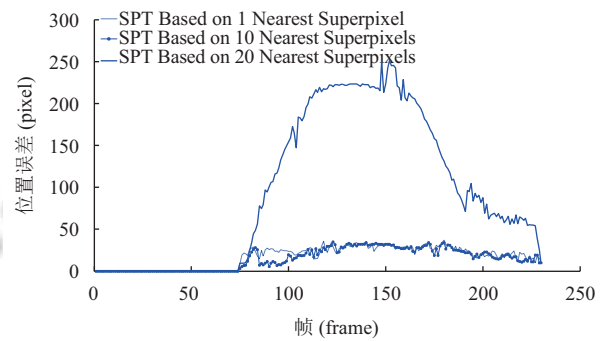


图 6 不同近邻个超像素的 SPT 的位置误差

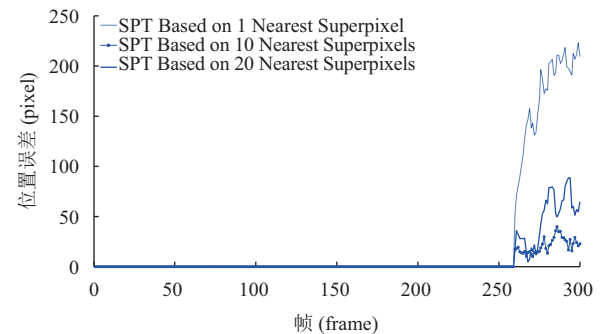


图 7 不同近邻个超像素的 SPT 的位置误差

图6中,不同近邻个超像素的SPT在skating1数据集第75~230帧上位置误差情况,1、10、20近邻个超像素的SPT对应位置总误差3848、3439、22189像素。

图7中,不同近邻个超像素的SPT在第260~300帧上位置误差情况,1、10、20近邻个超像素的SPT对应位置总误差6952.6、862.7、1968.5像素;10近邻个超像素(*型实线)的SPT表现最好,其他近邻的SPT总误差相对较大。

图8的Lemming数据集中,存在光照变化、尺度变化、不同程度遮挡、快速运动、离面旋转等情况。在这种背景下,基于历史外观模型的IVT由于外观模型不能很好的应对外观变化,所以在目标姿态和尺度都变化不大时表现良好,而在姿态和尺度发生较大变化时表现不佳。基于边信息的TLD由于未使用基于中层特征的外观模型,所以在跟踪姿态、尺度发生较大变化的目标时效果不如KSPT算法。在跟踪目标发生遮挡的第332帧、快速运动的第399帧、离面旋转的第997、1299帧,KSPT均表现出相当的优异性。



图8 KSPT(细线)、TLD(粗线)、IVT(虚线)在Lemming数据集第332、399、997、1299帧上的跟踪结果。



图9 KSPT(细线)、TLD(粗线)、IVT(虚线)在bolt数据集第20、92、222、350帧上的跟踪结果。

图9的Bolt数据集中,存在不同程度遮挡、变形、平面内运动、离面运动等情况。在这种背景下,基于边信息的TLD、基于历史外观模型的IVT的跟踪效果却差了很多,再次印证KSPT依托的中层特征外观模型的优越性。在跟踪目标因为运动引起形变的第20、92、222、350帧,KSPT均表现出相当的优异性。

4 结语

针对现有超像素跟踪算法仅依据特征池中最近邻超像素划定当前超像素簇归属造成分类错误的问题,本文提出一种健壮的超像素跟踪算法。该算法通过历史样本的平均中心误差,确定最佳近邻个数。根据最佳近邻个超像素划定当前超像素的簇归属,减弱了部分遮挡情况下特征置信度对候选样本的干扰,从而有效提升中心定位成功率和成功跟踪到的帧数。考虑到均值漂移聚类算法的耗时性,正在探索其他有效、快速的算法;运动模型中的速度、尺度参数都是经验值,不能实现自适应变化,尤其是出现撑杆跳等快速运动情况就很难发挥作用。另外,鉴于简单HIS颜色特征尚且取得不错的效果,相信更好特征的出现,一定能够很好的提升跟踪效果。

参考文献

- 1 Yilmaz A, Javed O, Shah M. Object tracking: A survey. ACM Computing Surveys, 2006, 38(4): 13. [doi: 10.1145/1177352]
- 2 Salari V, Sethi IK. Feature point correspondence in the presence of occlusion. IEEE Trans. on Pattern Analysis and Machine Intelligence, 1990, 12(1): 87-91. [doi: 10.1109/34.41387]
- 3 Veenman CJ, Reinders MJT, Backer E. Resolving motion correspondence for densely moving points. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2001, 23(1): 54-72. [doi: 10.1109/34.899946]
- 4 Broida TJ, Chellappa R. Estimation of object motion parameters from noisy images. IEEE Trans. on Pattern Analysis and Machine Intelligence, 1986, 8(1): 90-99.
- 5 Bar-Shalom Y, Fortmann TE. Tracking and data association. San Diego, CA, USA: Academic Press Professional, Inc., 1988.
- 6 Streit RL, Luginbuhl TE. Maximum likelihood training of probabilistic neural networks. IEEE Trans. on Neural Networks, 1994, 5(5): 764-783. [doi: 10.1109/72.317728]
- 7 Comaniciu D, Ramesh V, Meer P. Kernel-based object trac-

- king. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2003, 25(5): 564–577. [doi: [10.1109/TPAMI.2003.1195991](https://doi.org/10.1109/TPAMI.2003.1195991)]
- 8 Shi JB, Tomasi. Good features to track. 1994 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1994. Seattle, WA, USA. 1994. 593–600.
- 9 Tao H, Sathyanarayana HS, Kumar R. Object tracking with Bayesian estimation of dynamic layer representations. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2002, 24(1): 75–89. [doi: [10.1109/34.982885](https://doi.org/10.1109/34.982885)]
- 10 Black MJ, Jepson AD. EigenTracking: Robust matching and tracking of articulated objects using a view-based representation. International Journal of Computer Vision, 1998, 26(1): 63–84. [doi: [10.1023/A:1007939232436](https://doi.org/10.1023/A:1007939232436)]
- 11 Avidan S. Support vector tracking. Proc. of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001. Kauai, HI, USA. 2001. 184–191.
- 12 Isard M, Blake A. CONDENSATION-Conditional density propagation for visual tracking. International Journal of Computer Vision, 1998, 29(1): 5–28. [doi: [10.1023/A:1008078328650](https://doi.org/10.1023/A:1008078328650)]
- 13 Bertalmio M, Sapiro G, Randall G. Morphing active contours: A geometric approach to topology-independent image segmentation and tracking. Proc. of 1998 International Conference on Image Processing. Chicago, IL, USA. 1998. 318–322.
- 14 Kang JM, Cohen I, Medioni G. Object reacquisition using invariant appearance model. Proc. of the 17th International Conference on Pattern Recognition. Cambridge, UK. 2004. 759–762.
- 15 Jepson AD, Fleet DJ, El-Maraghi TF. Robust online appearance models for visual tracking. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2003, 25(10): 1296–1311. [doi: [10.1109/TPAMI.2003.1233903](https://doi.org/10.1109/TPAMI.2003.1233903)]
- 16 Lim J, Ross DA, Lin RS, *et al.* Incremental learning for visual tracking. Advances in Neural Information Processing Systems. Vancouver, Canada. 2004. 793–800.
- 17 Collins RT, Liu YX, Leordeanu M. Online selection of discriminative tracking features. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2005, 27(10): 1631–1643. [doi: [10.1109/TPAMI.2005.205](https://doi.org/10.1109/TPAMI.2005.205)]
- 18 Adam A, Rivlin E, Shimshoni I. Robust fragments-based tracking using the integral histogram. 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. New York, NY, USA. 2006. 798–805.
- 19 Grabner H, Leistner C, Bischof H. Semi-supervised on-line boosting for robust tracking. Proc. of the 10th European Conference on Computer Vision: Part I. Marseille, France. 2008. 234–247.
- 20 Kwon J, Lee KM. Tracking of a non-rigid object via patch-based dynamic appearance modeling and adaptive basin hopping monte carlo sampling. IEEE Conference on Computer Vision and Pattern Recognition(CVPR 2009). Miami, FL, USA. 2009. 1208–1215.
- 21 Kwon J, Lee KM. Visual tracking decomposition. 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2010). San Francisco, CA, USA. 2010. 1269–1276.
- 22 Ren XF, Malik J. Tracking as repeated figure/ground segmentation. IEEE Conference on Computer Vision and Pattern Recognition. Minneapolis, MN, USA. 2007. 1–8.
- 23 Yang F, Lu HC, Yang MH. Robust superpixel tracking. IEEE Trans. on Image Processing, 2014, 23(4): 1639–1651. [doi: [10.1109/TIP.2014.2300823](https://doi.org/10.1109/TIP.2014.2300823)]
- 24 陈桂景, 王尧弘. 自适应近邻判别分析. 数学物理学报, 1996, 16(S1): 9–19.
- 25 Ross DA, Lim J, Lin RS, *et al.* Incremental learning for robust visual tracking. International Journal of Computer Vision, 2008, 77(1-3): 125–141. [doi: [10.1007/s11263-007-0075-7](https://doi.org/10.1007/s11263-007-0075-7)]
- 26 Kwon J, Lee KM. Visual tracking decomposition. 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2010). San Francisco, CA, USA. 2010. 1269–1276.
- 27 Bao CL, Wu Y, Ling HB, *et al.* Real time robust L1 tracker using accelerated proximal gradient approach. IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2012). Providence, RI, USA. 2012. 1830–1837.
- 28 Kalal Z, Matas J, Mikolajczyk K. P-N learning: Bootstrapping binary classifiers by structural constraints. 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2010). San Francisco, CA, USA. 2010. 49–56.
- 29 Hare S, Golodetz S, Saffari A, *et al.* Struck: Structured output tracking with kernels. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2016, 38(10): 2096–2109. [doi: [10.1109/TPAMI.2015.2509974](https://doi.org/10.1109/TPAMI.2015.2509974)]