

基于视觉显著性特征的粒子滤波跟踪算法

吴世东, 鲍华, 张陈斌, 陈宗海

(中国科学技术大学自动化系, 安徽合肥 230027)

摘要:针对复杂环境下的运动目标跟踪问题,提出了一种基于视觉显著性特征的粒子滤波跟踪算法.该算法利用显著性检测算法对序列图片进行检测,生成视觉显著图,然后利用二阶自回归模型对目标状态进行预测,再根据中心强化-四周弱化的机制,生成最终显著图.利用视觉显著图中目标区域像素值较大的特点,提取视觉显著性特征,与颜色特征进行自适应融合,从而完成跟踪.实验结果表明,该算法能够有效应对跟踪过程中出现的场景光照变化和目标姿态变化等问题,具有较强的鲁棒性.

关键词:目标跟踪;视觉显著图;粒子滤波;显著性特征

中图分类号:TP391 **文献标识码:**A **doi:**10.3969/j.issn.0253-2778.2015.11.009

引用格式: Wu Shidong, Bao Hua, Zhang Chenbin, et al. Particle filter tracking based on visual saliency feature [J]. Journal of University of Science and Technology of China, 2015, 45(11): 934-942.

吴世东, 鲍华, 张陈斌, 等. 基于视觉显著性特征的粒子滤波跟踪算法[J]. 中国科学技术大学学报, 2015, 45(11): 934-942.

Particle filter tracking based on visual saliency feature

WU Shidong, BAO Hua, ZHANG Chenbin, CHEN Zonghai

(Department of Automation, University of Science and Technology of China, Hefei 230027, China)

Abstract: To address the problem of moving object tracking in complicated scenes, a particle filter tracking algorithm based on visual saliency feature was presented. The algorithm detects the object in the image with the saliency detection algorithm to get saliency maps. Target states are predicted using the second-order autoregressive model, and the final saliency map is obtained with the center-strengthening and edge-weakening mechanism. The saliency feature is extracted according to the phenomenon that in the saliency map pixel value is greater when the pixel is in the target area, and is then fused with the color feature adaptively to complete tracking. Experimental results show that the algorithm can effectively deal with the situation when illumination and appearance change.

Key words: visual tracking; saliency map; particle filter; saliency feature

收稿日期:2015-04-07;修回日期:2015-11-05

基金项目:国家自然科学基金(61375079)资助.

作者简介:吴世东,男,1989年生,硕士生.研究方向:计算机视觉与模式识别. E-mail:wsd123@mail.ustc.edu.cn

通讯作者:陈宗海,教授. E-mail:chenzh@ustc.edu.cn

0 引言

运动目标跟踪是计算机视觉领域中一个重要的分支,其相关研究成果在视频监控、图像检索、增强现实、人机交互和犯罪行为分析等方面均有广泛的应用^[1-2]. 在目标跟踪过程当中,场景光照变化、目标局部或全局遮挡、目标姿态变化以及目标发生突变运动等情况,都对跟踪鲁棒性造成了极大的挑战. 为了在复杂环境下实现鲁棒跟踪,研究人员根据不同的跟踪环境,提出了许多有针对性的跟踪算法. 其中,粒子滤波算法^[3-4]是基于蒙特卡罗模拟的采样跟踪算法,由于具有解决非线性、非高斯问题的优势,在目标跟踪当中有着广泛的应用. 在粒子滤波跟踪算法当中,常用的特征有颜色、方向、纹理等,其中,颜色特征对于目标姿态变化和部分遮挡不敏感,且计算简单,是最常使用的特征.

最近几年,视觉显著性^[5]获得了广泛关注,出现了很多视觉显著性检测算法^[6-9]. 视觉显著性机制就是利用人类视觉中的选择性注意机制,从海量的视觉信息当中筛选出有用的信息,提高图像处理效率. 随着视觉显著性检测算法的日趋成熟,视觉显著性在目标检测^[10]、目标识别^[11]等领域都有着广泛的应用. 最近几年,视觉显著性在目标跟踪当中的应用也多了起来. 朱明清等^[12]提取序列图像颜色直方图,由此生成视觉显著图,利用显著图中目标区域显著度较高的特点实现了对丢失目标的再跟踪;赵宇宙等^[13]模拟人类视觉中具有的选择性注意机制,提出一种在线实时选择目标区域显著子块的跟踪方法,该方法提取目标区域中与背景差异较小的小块用于跟踪,利用这些小块与跟踪目标之间的空间关系得到跟踪结果,该方法提高了应对部分遮挡的能力;江晓莲等^[14]针对突变目标的跟踪,提出一种基于视觉显著图的蒙特卡罗跟踪算法,该算法首先建立目标的运动状态模型,由此判定目标运动是否发生突变,若目标发生突变运动,则采用生成的视觉显著图来提取目标显著性信息,由此来引导马尔可夫链的构造,提高目标区域粒子接受概率,从而完成对突变目标的跟踪;Yan 等^[15]将目标跟踪过程当成是目标与背景的持续分离问题,提出一种利用视觉显著性的在线跟踪算法,为解决分类器当中普遍存在的正负样本准确性不高的问题,文章利用生成的显著图来指导正样本的选择,有效地提高了样本的质量,确保了跟踪的精度;Zhang 等^[16]提出一种基于

环境感知显著性的关键目标识别与跟踪算法,文章首先通过将目标与其周围环境进行交互感知提高视觉显著图的精度,然后利用得到的精细显著图来引导关键目标的识别与跟踪,在跟踪鲁棒性与实时性方面都达到了较好的效果.

通过分析以上跟踪算法可以发现,当前在目标跟踪算法当中,对视觉显著性信息的利用仅限于用显著图来指导或引导跟踪,并没有将显著性当成一种特征来对待,这导致这些算法对目标显著性信息利用不充分、不完全. 有鉴于此,本文提出一种颜色特征与视觉显著性特征相结合的运动目标跟踪算法,在该算法当中,视觉显著性特征能够有效弥补颜色特征对于场景光照变化敏感的问题,且当目标发生姿态变化时,显著图能够进行目标位置的大致定位,避免造成跟踪漂移.

1 粒子滤波跟踪框架

粒子滤波算法的核心思想是贝叶斯滤波理论和蒙特卡罗方法. 具体做法是根据 t 时刻系统状态的观测结果 $\mathbf{Z}_t = \{\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_t\}$, 利用贝叶斯概率理论找出系统的后验概率分布 $P(\mathbf{x}_t | \mathbf{z}_{1:t})$, 再根据蒙特卡罗方法,用有限样本点状态值的加权求和运算得到跟踪结果,即计算出系统的状态 \mathbf{x}_t .

假设系统的状态模型和观测模型为

$$\left. \begin{aligned} \mathbf{x}_t &= f(\mathbf{x}_{t-1}, \mathbf{u}_t) \\ \mathbf{z}_t &= h(\mathbf{x}_t, \mathbf{v}_t) \end{aligned} \right\} \quad (1)$$

式中, \mathbf{x}_t 为目标在 t 时刻的状态,一般表示目标的位置和尺度信息; \mathbf{z}_t 为 t 时刻对系统状态的观测结果,实际表示为对图像信息的采样观测结果; f 和 h 分别表示系统状态函数和观测函数; \mathbf{u}_t 和 \mathbf{v}_t 分别为系统噪声和观测噪声.

贝叶斯滤波过程分为状态预测和状态更新两个步骤. 预测过程是指在不清楚 t 时刻系统状态观测结果 \mathbf{z}_t 的情况下,通过 $t-1$ 时刻的概率密度函数 $P(\mathbf{x}_{t-1} | \mathbf{z}_{1:t-1})$ 和系统状态转移模型 $P(\mathbf{x}_t | \mathbf{x}_{t-1})$ 来近似得到 t 时刻系统后验概率密度函数 $P(\mathbf{x}_t | \mathbf{z}_{1:t-1})$, 具体公式如下:

$$P(\mathbf{x}_t | \mathbf{z}_{1:t-1}) = \int P(\mathbf{x}_t | \mathbf{x}_{t-1}) P(\mathbf{x}_{t-1} | \mathbf{z}_{1:t-1}) d\mathbf{x}_{t-1} \quad (2)$$

更新过程是指利用 t 时刻的观测结果 \mathbf{z}_t 来对系统进行修正,根据贝叶斯公式, t 时刻系统后验概率密度函数 $P(\mathbf{x}_t | \mathbf{z}_{1:t})$ 为

$$P(\mathbf{x}_t | \mathbf{z}_{1:t}) = \frac{P(\mathbf{z}_t | \mathbf{x}_t)P(\mathbf{x}_t | \mathbf{z}_{1:t-1})}{P(\mathbf{z}_t | \mathbf{z}_{1:t-1})} \quad (3)$$

式中, $P(\mathbf{z}_t | \mathbf{x}_t)$ 称为似然性, 实为系统状态观测函数.

用一组带有权重的粒子来近似表示后验概率密度函数 $P(\mathbf{x}_t | \mathbf{z}_{1:t})$, 设带权重的粒子记为 $\{\mathbf{x}_t^i, \omega_t^i\}_{i=1}^N$, 其中 \mathbf{x}_t^i 为样本中第 i 个粒子的状态, ω_t^i 为其相应权重, $\sum_{i=1}^N \omega_t^i = 1$, N 为粒子的个数. 根据蒙特卡罗原理, 可以用一个离散的加权式来逼近 t 时刻的后验概率密度, 具体公式为

$$P(\mathbf{x}_{0:t} | \mathbf{z}_{1:t}) \approx \sum_{i=1}^N \omega_t^i \delta(\mathbf{x}_{0:t} - \mathbf{x}_{0:t}^i) \quad (4)$$

式中, 粒子权重 ω_t^i 与粒子观测概率相关, $\delta(\cdot)$ 为 Dirac 函数.

这样, 系统状态 \mathbf{x}_t 的滤波估计结果可以用粒子状态值加权求和表示, 其计算公式为

$$\hat{\mathbf{x}}_t = \sum_{i=1}^N \omega_t^i \mathbf{x}_t^i \quad (5)$$

式中, $\hat{\mathbf{x}}_t$ 即为目标状态滤波结果, 也即最终跟踪结果.

为防止粒子退化, 采用重要性重采样^[17]的方法对粒子进行重采样, 其核心思想是复制权重较高的粒子, 抛弃权重较小的粒子, 产生新的粒子集.

2 视觉显著性检测

视觉注意机制是指模拟生物视觉, 特别是人类视觉中的关注机制, 该机制能够帮助人们从海量的图像信息中寻找出感兴趣的目标, 并迅速将注意焦点放到该目标当中, 忽略目标周围复杂背景的影响, 从而提高图像信息处理效率.

FT 算法^[7]是近年出现的一种基于频域调制的显著性检测算法, 该算法具有较好的检测效果, 得到的显著图中目标区域边界保存良好, 且计算量小, 在目标检测和图像标注方面都有着广泛的应用.

2.1 FT 算法计算模型

FT 算法计算非常简单. 对于一幅图像 I , 设其宽、高分别为 W 和 H , 则像素点 (m, n) 的显著性值为

$$S(m, n) = |I_\mu - I_{Whc}| \quad (6)$$

式中, I_μ 为图像 I 的像素值均值, I_{Whc} 为图像 I 的高斯模糊结果.

为提高颜色特征对目标与背景的区别性, 整个特征提取与计算是在 Lab 空间中完成的, 这样式

(6) 可以重新定义为

$$S(m, n) = \|I_\mu - I_{Whc}(m, n)\| \quad (7)$$

式中, I_μ 为图像 I 在 Lab 空间中的特征向量均值, $I_{Whc}(m, n)$ 为对像素点 (m, n) 处像素值进行高斯模糊化的结果(高斯滤波公式采用 3×3 或 5×5 形式的可分离二项式内核, 所谓可分离是指对水平和垂直方向分别进行高斯滤波处理), I_μ 与 $I_{Whc}(m, n)$ 在 Lab 空间中的形式为 $[L, a, b]^T$, $\|\cdot\|$ 为 L_2 范数, 在此处采用欧氏距离进行向量间距离的计算.

2.2 显著性增强算法

在实际的跟踪过程当中, 由于复杂场景的影响, 目标区域在生成的显著图中可能不是唯一显著的区域, 从而会影响生成的显著图的质量. 例如, 当场景图片中出现与目标颜色相似的物体时, 在生成的显著图中, 目标与该物体可能都会显著; 当场景中出现强光时, FT 算法得到的显著图中强光区域显著度较高, 间接影响目标显著性. 为此, 本文根据中心强化-四周弱化机制, 对 FT 算法生成的显著图中非目标区域显著性进行抑制操作, 从而降低非目标区域在显著图上的显著度, 突出目标区域的显著性.

根据目标运动平滑的假设, 目标相邻几帧在图像中的位置比较接近, 因此可根据 t 时刻之前两帧对目标的跟踪结果 \mathbf{x}_{t-1} 与 \mathbf{x}_{t-2} , 并利用二阶自回归模型预测 t 时刻目标的位置:

$$\mathbf{x}_t = 2\mathbf{x}_{t-1} - \mathbf{x}_{t-2} + \mathbf{v}_t \quad (8)$$

式中, $\mathbf{x}_t = [x, y, s]$ 为目标在 t 时刻的预测状态, (x, y) 为目标中心位置坐标, s 为目标相对于第一帧的尺度缩放大小; \mathbf{v}_t 为零均值高斯模型.

对于图像 I , 利用 FT 模型计算得到显著图 S 之后, 为突出目标区域的显著度, 提高显著图的质量, 根据中心强化-四周弱化机制, 定义目标区域显著性增强图 SAL 为

$$SAL(m, n) = \frac{\sqrt{W^2 + H^2} - \sqrt{(m-x)^2 + (n-y)^2}}{\sqrt{W^2 + H^2}} S(m, n) \quad (9)$$

图 1 是后面跟踪实验用到的两组测试序列当中两张图片的显著性检测结果. 从图中可以看到, 第 1 行中, 目标在空中高速运动, 伴随着尺度和姿态变化, 但是在生成的显著图中, 目标区域依然显著; 第 2 行中, 场景发生剧烈的光照变化, FT 检测结果得到的显著图中, 除目标区域外, 强光等背景区域显著度也较高, 但是在用显著性增强算法处理之后, 背景

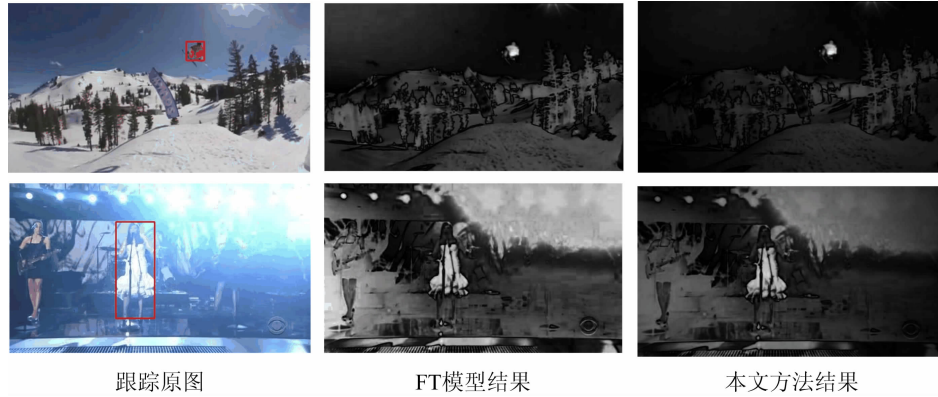


图 1 显著性检测结果

Fig. 1 Results of saliency detection

区域的显著度被极大地抑制,突出了目标区域的显著性.

2.3 显著性特征的定义与提取

所谓视觉显著性特征,是一种能直观表现出目标区域与背景区域差异的特征,具体在显著图中体现为目标区域显著度较高,而背景区域显著度较低.显著性特征与颜色、纹理等特征不同,不需要建立特征模板,避免了跟踪过程中经常出现的模板漂移现象.

显著性检测的结果是得到一幅单通道显著图 SAL,在 SAL 中,目标区域像素值较大,而背景区域像素值较小,因此可以统计每个粒子区域内的像素值总和,以此来度量粒子的权重大小.

按照上述理论, t 时刻样本中的第 i 个粒子的显著性特征观测概率定义为

$$P(z_i^s | x_i^t) = \frac{1}{S_t} \sum_{(m,n) \in S_i^t} SAL(m,n) \quad (10)$$

式中, S_t 为相关参数, S_i^t 为第 i 个粒子在图像中的区域范围.

3 目标跟踪框架

3.1 特征融合机制

采用颜色特征和视觉显著性特征一起来进行跟踪,颜色特征的提取是在接近人类视觉习惯的 HSV 颜色空间中完成的.采用 HSV 颜色直方图来表示目标的颜色特征表现模型,令 $\mathbf{H}_c = \{q^u\}_{u=1}^{\Gamma}$ 为目标颜色特征模板, $\mathbf{H}_i = \{p_i^u\}_{u=1}^{\Gamma}$ 为在第 i 个粒子区域范围内提取的颜色直方图,其中 Γ 为对像素值进行量化的区间总个数, q^u 为落在第 u 个量化区间内的像素值个数占整个图片像素值个数的比例.这样 t 时

刻样本中的第 i 个粒子的颜色特征观测概率定义为

$$P(z_i^c | x_i^t) = \rho(\mathbf{H}_i^t, \mathbf{H}_c) \quad (11)$$

式中, $\rho(\mathbf{H}_i^t, \mathbf{H}_c) = \sum_{u=1}^{\Gamma} \sqrt{q^u p_i^u}$ 为直方图间相似度计算公式.

特征融合方法一般有乘性机制和加性机制两种,所谓乘性机制就是将各特征的观测概率相乘,加性机制就是将它们相加.根据它们的定义容易得出,乘性机制融合方法能够提高粒子重要性区分度,却容易放大噪声,而加性机制融合方法能够降低噪声干扰,却不利于提高粒子重要性区分度.本文将它们两者的优势结合起来,定义特征融合公式为

$$P_1(z_i | x_i^t) = [\alpha + \gamma P(z_i^c | x_i^t)] [\alpha + (1 - \gamma) P(z_i^s | x_i^t)] \quad (12)$$

式中, α 为 0~1 之间的常数; γ 为颜色特征的重要性参数,设粒子个数为 N ,令

$$\gamma = \frac{1}{N} \sum_{i=1}^N P(z_i^c | x_i^t).$$

式(12)展开之后会出现前面所说的两种特征观测概率之间的乘法和加法运算,从而充分体现乘性和加性融合机制各自的优势.参数 γ 实为各个粒子颜色特征观测概率均值,该值反映了复杂环境对颜色特征的干扰程度.

当式(12)中参数 α 取得较大时,可能会出现融合观测概率大于 1 的情况,因此对 $P_1(z_i | x_i^t)$ 进行归一化操作,得到最终观测概率 $P(z_i | x_i^t)$:

$$P(z_i | x_i^t) = \frac{1}{\sum_{i=1}^N P_1(z_i | x_i^t)} P_1(z_i | x_i^t) \quad (13)$$

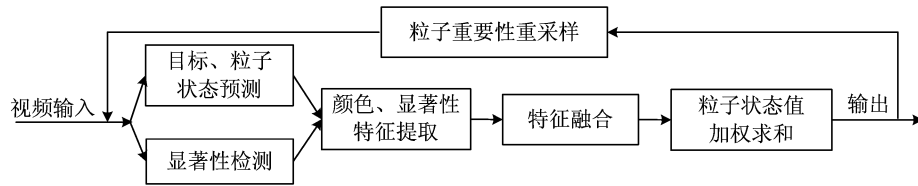


图 2 基于视觉显著性特征的粒子滤波跟踪流程

Fig. 2 Process of particle filter tracking based on visual saliency feature

3.2 目标跟踪算法

第一帧在图像中手动选取要跟踪的目标,并提取目标区域 HSV 颜色直方图作为颜色特征的模板,从第二帧开始,首先利用二阶自回归模型对目标和粒子状态分别进行预测,并同时图像进行显著性检测,生成显著图,然后分别提取各粒子颜色特征和视觉显著性特征,得到它们的观测概率,再进行特征融合,最后对粒子状态值加权求和得到跟踪结果,跟踪完成之后对粒子进行重采样. 整体跟踪流程如图 2 所示.

整体算法流程如下:

Step 1 $t=1$, 目标和粒子初始化. 因式(8)计算需要, 令 $\mathbf{x}_0 = \mathbf{x}_1, \mathbf{x}_0^i = \mathbf{x}_1^i, i=1, 2, \dots, N$;

Step 2 状态预测和显著性检测:

① 令 $t=t+1$;

② 利用式(8)进行目标和粒子状态预测;

③ 利用式(7), (9)对原始图像进行显著性检测, 得到显著性增强图 SAL;

Step 3 利用式(10), (11)分别得到视觉显著性特征和颜色特征观测概率 $P(\mathbf{z}_t^i | \mathbf{x}_t^i)$ 和 $P(\mathbf{z}_t^c | \mathbf{x}_t^i)$;

Step 4 特征融合和粒子权重计算:

① 利用式(12)进行特征融合, 并用式(13)进行归一化处理, 得到融合后概率 $P(\mathbf{z}_t | \mathbf{x}_t^i)$;

② 粒子权重计算公式为

$$w_t^i = e^{-\beta(1-P(\mathbf{z}_t | \mathbf{x}_t^i))} \quad (14)$$

式中, β 为相关参数.

粒子权重归一化公式为

$$\hat{w}_t^i = w_t^i / \sum_{i=1}^N w_t^i \quad (15)$$

Step 5 得到跟踪结果 $\hat{\mathbf{x}}_t$,

$$\hat{\mathbf{x}}_t = \sum_{i=1}^N \hat{w}_t^i \mathbf{x}_t^i \quad (16)$$

Step 6 采用文献[17]方法进行粒子重要性重采样;

Step 7 是否是最后一帧, 否则返回 Step 2.

4 仿真实验和结果分析

为验证本文算法的跟踪鲁棒性, 选取了 4 个具有针对性的视频序列进行测试. 重点考察了场景光照变化和目標姿态变化情况下的跟踪结果, 以及分析在跟踪过程中视觉显著性特征所起的作用, 并将其与一些经典算法进行了对比实验, 如 PF^[3], ST^[14], RT^[15], FT^[18] 和 CT^[19].

4.1 实验设置

算法是在 C++ 环境下实现的, 运行于 Intel Core 2 Duo 6600 机器上. 实验参数设置为: 式(7)中高斯模糊化公式采用 3×3 形式的可分离二项式内核, 式(10)中 $S_t = \sum_{(m,n) \in SAL} SAL(m,n)$, 式(12)中 $\alpha=0.5$, 式(14)中 $\beta=20$, HSV 颜色直方图中 bin 值(量化通道个数)设为 $16 \times 16 \times 16$, 粒子个数 $N=100$.

4.2 结果与分析

实验结果如图 3 所示, 各算法采用了不同的颜色进行标注, 实验中各被比较算法采用与作者原算法相同的参数设置, 各个序列的 ground truth 由视频的提供者给出.

实验 1 对发生光照变化和尺度变化的目标进行跟踪. Singer 序列包含了场景光照变化和目標尺度变化, 各算法表现有优有劣. 其中, 在第 60~130 帧之间, 场景发生了剧烈的光照变化. 基于压缩感知的 CT 算法由于使用强度特征训练分类器, 而颜色特征对光照变化敏感, 从而无法在光照变化情况下对目标和背景进行正确分类, 跟踪效果不佳; 基于固定分块跟踪的 FT 算法也是采用强度特征对目标进行跟踪, 因此也无法在光照变化情况下保证跟踪精度; 采用单一颜色特征的 PF 算法跟上述两种算法一样, 无法正确跟踪上目标; 基于视觉显著性的 ST 算法得到的显著图中, 强光区域显著度也较高, 因而无法实现对丢失目标的再跟踪, 出现了跟踪漂移现象; 基于视觉显著性的 RT 算法能够利用显著图来

引导分类器中正样本的选择,提高了样本精度,跟踪结果取得了次优的结果. 而本文算法除采用颜色特征外,还引入视觉显著性特征对目标进行跟踪. 从图 3 中可以看出,在光照变化情况下,目标区域在显著图中依然显著,此时显著性特征可以被视作一种光照不变特征,且本文选用的特征融合策略使得在光照变化环境下,显著性特征在特征融合过程中起主要作用,极大地提高了跟踪精度.

实验 2 对发生姿态变化的目标进行跟踪. Tiger 序列包含了目标姿态变化. 从实验结果来看, CT 算法和本文算法取得了不错的跟踪精度. CT 算法由于采用压缩感知的方法对图像特征进行降维,且每帧都通过在线学习的方法更新分类器,因此对姿态变化适应性强,取得了最好的跟踪效果;本文算法采用的视觉显著性特征受到姿态变化影响较小(从图 3 可以看出),取得了次优的结果;RT 与 ST

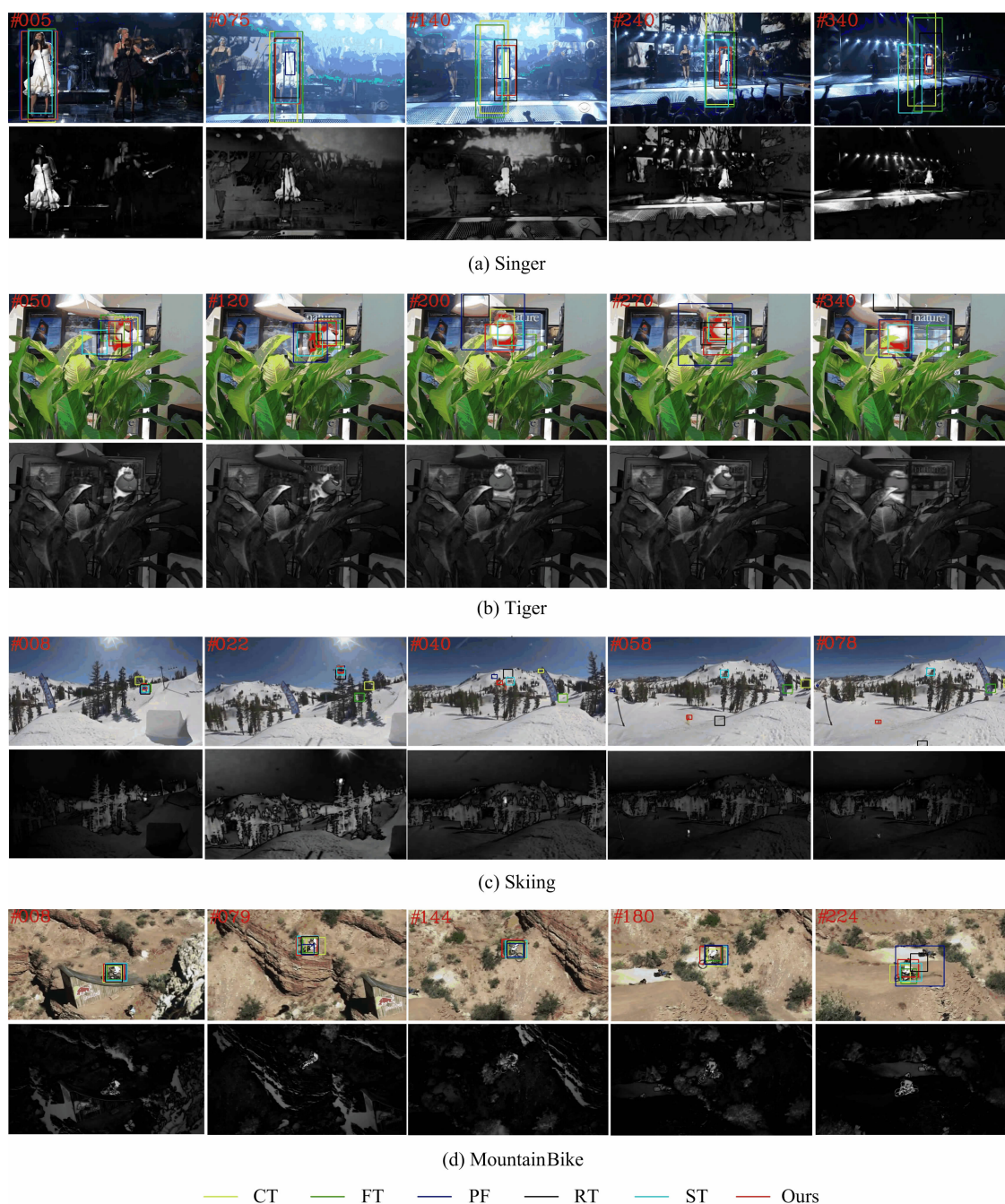


图 3 各跟踪算法的跟踪结果以及各序列显著性检测结果

Fig. 3 Tracking results of each algorithm and saliency detection results of each sequence

算法在一定程度上利用了视觉显著性信息用于跟踪,但是由于对显著性信息的利用不充分,跟踪结果一般;PF 算法在跟踪过程中一度出现了跟踪漂移,但是颜色特征对姿态变化具有一定抵抗性,最后恢复了对目标的跟踪;FT 算法采用固定分块的策略无法适用目标姿态变化,出现了跟踪漂移现象。

实验 3 对发生姿态变化和尺度变化的目标进行跟踪。Skiing 序列包含了目标姿态变化和尺度变化,除本文算法外,其余跟踪算法都出现了跟踪漂移。其中,FT 和 CT 算法采用的强度特征对目标和背景的区别性不强,且它们在跟踪过程当中无法进行尺度的自适应调整,导致了跟踪漂移(第 12 帧);PF 算法采用的 HSV 直方图对目标姿态变化具有一定的抵抗性,刚开始可以跟踪上目标,但是当目标姿态变化和尺度变化幅度较大时,颜色特征模板出现了漂移,导致跟踪失败(第 50 帧);RT 与 ST 算法在跟踪过程中虽然避免了跟踪结果的急剧漂移,但是由于算法本身对目标姿态变化和尺度变化适应性不强,最终也出现了跟踪失败;本文算法使用的颜色特征和视觉显著性特征受到目标姿态变化和尺度变

化影响较小。

实验 4 对发生姿态变化的目标进行跟踪。MountainBike 序列包含了目标姿态变化,各算法虽然都未出现明显的跟踪漂移,但是在跟踪精度上有一定的差异,本文算法跟踪精度最高。其中,在第 180~220 帧之间,目标发生姿态变化的程度较大,且伴随着一定程度的快速运动,除本文算法外,其他跟踪算法都出现了较为明显的跟踪偏差(见图 4)。

为了更加直观地表示各跟踪算法的跟踪精度,定义中心位置误差 C_t 为

$$C_t = \sqrt{(x - x_g)^2 + (y - y_g)^2} \quad (17)$$

式中, (x, y) 为跟踪中得到的目标状态 x_t 的中心位置坐标, (x_g, y_g) 为 t 时刻 ground truth 中目标的中心位置坐标。

图 4 为不同视频序列中各个算法的中心位置误差曲线。由图可以看出本文算法在各个视频序列中的跟踪误差曲线比较平稳,一直稳定在一个较小的误差范围内,体现出本文算法的稳定性,能够实现对照光变化、尺度变化和剧烈表观变化的目标的稳定跟踪。表 1 给出的是不同视频序列的各算法的平均

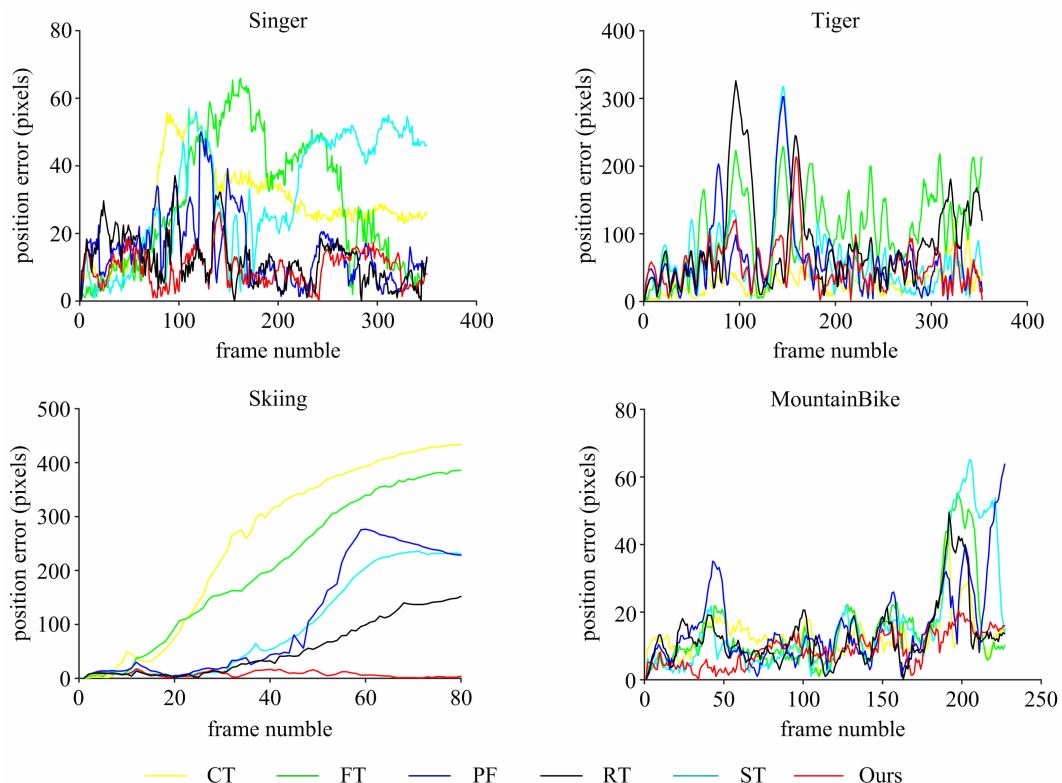


图 4 各跟踪算法的中心位置误差

Fig. 4 The center position error of each tracking algorithm

中心位置误差,红色、蓝色分别表示最优、次优结果.由表 1 可以看出,本文算法对光照发生剧烈变化的 Singer 序列、表观和尺度发生剧烈变化的 Skiing 序列、目标姿态发生变化的 MountainBike 序列取得了最优结果,对于表观发生变化的 Tiger 序列取得了次优的结果.

从这几组实验结果可以看出,相对于采用单一颜色特征的 CT, FT, PF 算法和利用视觉显著性来引导跟踪的 ST, RT 算法而言,本文将显著性当成一种特征来对待,提出的视觉显著性特征能够通过颜色特征进行自适应融合权重的方式弥补颜色特征的不足,提高跟踪精度.在 Singer 序列中,光照发生剧烈变化,在 Skiing 序列中,目标表观、尺度发生了剧烈变换,这给跟踪带来了极大的难度,但是,本文算法依然取得了较好的跟踪效果.总结来说,本文算法能够取得较好跟踪结果主要得益于两点:①本文提出的视觉显著性特征对图像显著性信息的利用较为充分、完全,使其在一些跟踪序列中对光照变化和姿态变化等干扰因素具有不变性,即这些干扰因素对目标区域在显著图中的显著性影响较小;②由于在跟踪过程中目标真实位置的未知性,一般采用的建立目标表观模型和模板更新的跟踪算法容易出现模板漂移,而本文的显著性特征并不需要建立模板,因此即使某几帧出现了较大的跟踪偏差,也具有恢复跟踪的可能性,不会出现一般的漂移现象.

表 1 平均中心位置误差

Tab. 1 The average center position error	CT	FT	PF	RT	ST	Ours
Singer	27.2	32.4	15.3	12.1	29.3	10.2
Tiger	27.5	100.2	50.5	34.2	48.2	32.5
Skiing	230.3	200.4	100.3	50.2	89.4	15.5
MountainBike	17.2	18.1	19.6	18.8	21.1	10.6

5 结论

针对跟踪过程当中出现的光照变化、姿态变化、尺度变化等问题,本文在图像序列的显著图中提取视觉显著性特征,与颜色特征进行融合跟踪.实验结果显示,本文提出的视觉显著性特征和特征融合策略对视觉显著性信息进行了充分的利用,能够弥补单一颜色特征在跟踪过程中的不足,取得了较好的跟踪效果.本文方法的不足之处是,在一些测试序列当中,目标区域在显著图中显著度不高.考虑到

目标位置在第一帧是已知的,下一步拟从机器学习的角度,在显著性检测算法当中融入目标表观信息,构建“自顶向下”显著性检测算法,确保目标区域在显著图中的唯一显著性,从而应对跟踪过程中出现的遮挡、姿态变化等情况.

参考文献(References)

- [1] Li Xi, Hu Weiming, Shen Chunhua, et al. A survey of appearance models in visual object tracking[J]. ACM Transactions on Intelligent Systems and Technology, 2013, 4(4): 58-96.
- [2] Yilmaz A, Javed O, Shah M. Object tracking: A survey[J]. ACM Computing Surveys, 2006, 38(4): 1-45.
- [3] Pérez P, Hue C, Vermaak J, et al. Color-based probabilistic tracking [C]// Proceedings of the 7th European Conference on Computer Vision. Copenhagen: Springer Berlin Heidelberg, 2002: 661-675.
- [4] Zhu Mingqing, Wang Zhiling, Chen Zonghai. Visual tracking algorithm based on grey prediction model and particle filter[J]. Control and Decision, 2012, 27(1): 53-57.
朱明清, 王智灵, 陈宗海. 基于灰色预测模型和粒子滤波的视觉目标跟踪算法[J]. 控制与决策, 2012, 27(1): 53-57.
- [5] Borji A, Itti L. State-of-the-art in visual attention modeling [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(1): 185-207.
- [6] Itti L, Koch C, Niebur E. A model of saliency-based visual attention for rapid scene analysis [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(11): 1 254-1 259.
- [7] Achanta R, Hemami S, Estrada F, et al. Frequency-tuned salient region detection[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Miami, Florida, USA: IEEE Computer Society, 2009: 1 597-1 604.
- [8] Cheng M M, Zhang G X, Mitra N J, et al. Global contrast based salient region detection [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Colorado Springs, Colorado, USA: IEEE Computer Society, 2011: 409-416.
- [9] Qian Sheng, Lin Mingqiang, Chen Zonghai. Saliency detection based on conditional random field and image segmentation [J]. Acta Automatica Sinica, 2015, 41(4): 711-724.

- 钱生, 林明强, 陈宗海, 等. 基于条件随机场和图像分割的显著性检测[J]. 自动化学报, 2015, 41(4): 711-724.
- [10] Navalpakkam V, Itti L. An integrated model of top-down and bottom-up attention for optimizing detection speed[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York, NY USA: IEEE Computer Society, 2006: 2 049-2 056.
- [11] Gao D, Han S, Vasconcelos N. Discriminant saliency, the detection of suspicious coincidences, and applications to visual recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 31(6): 989-1 005.
- [12] Zhu Mingqing, Wang Zhiling, Chen Zonghai. Human visual intelligence and particle filter based robust object tracking algorithm[J]. Control and Decision, 2012, 27(11): 1 720-1 724.
朱明清, 王智灵, 陈宗海. 基于人类视觉智能和粒子滤波的鲁棒目标跟踪算法[J]. 控制与决策, 2012, 27(11): 1 720-1 724.
- [13] Zhao Yuzhou, Chen Zonghai. Robust object tracking by selecting local salient subregions [J]. Control and Decision, 2014, 29(10): 1 788-1 792.
赵宇宙, 陈宗海. 显著子区域在线选择的目标鲁棒跟踪[J]. 控制与决策, 2014, 29(10): 1 788-1 792.
- [14] Jiang Xiaolian, Li Cuihua, Li Xiongzong. 江晓莲, 李翠华, 李雄宗. Saliency based tracking method for abrupt motions via two-stage sampling[J]. Acta Automatica Sinica, 2014, 6(40): 1 098-1 107.
基于视觉显著性的两阶段采样突变目标跟踪算法[J]. 自动化学报, 2014, 6(40): 1 098-1 107.
- [15] Yan Jia, Chen Xi, Zhu Qiuping. Robust online tracking via adaptive samples selection with saliency detection[J]. EURASIP Journal on Advances in Signal Processing, 2013, 2013(1): 1-11.
- [16] Zhang Geng, Yuan Zejian, Zheng Nanning. Key object discovery and tracking based on context-aware saliency [J]. International Journal of Advanced Robotic Systems, 2013, 10(15): 1-12.
- [17] Liu J S. Metropolized independent sampling with comparisons to rejection sampling and importance sampling[J]. Statistics and Computing, 1996, 6(2): 113-119.
- [18] Adam A, Rivlin E, Shimshoni I. Robust fragments-based tracking using the integral histogram [C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. New York, USA: IEEE Computer Society, 2006: 798-805.
- [19] Hang K, Zhang L, Yang M H. Real-time compressive tracking[C]// Proceeding of 12th European Conference on Computer Vision. Florence: Springer Berlin Heidelberg, 2012: 864-877.