

特征点辅助的时空上下文目标跟踪与定位^{*}

翟敬梅, 刘 坤

(华南理工大学机械与汽车工程学院 广州 510640)

摘 要:针对动态目标跟踪中快速运动和目标遮挡而跟踪失败问题,提出了一种特征点辅助的时空上下文跟踪算法。首先提取目标特征点,通过特征点匹配和光流跟踪方法进行目标追踪,获得目标预估位置;其次,建立特征点变化率和时空上下文模型更新率关系模型,实时调控更新率,防止引入错误信息;最后,在预估位置区域内,构建局部上下文外观模型,计算与时空上下文模型的相关性获取置信图,进一步精确定位目标。算法在一组测试视频集中进行验证,相比目前4种主流算法(平均跟踪成功率最高为60%,平均跟踪误差最小为26.14 pixel),本算法综合性能达到最优,平均跟踪成功率为90%,平均跟踪误差为7.47 pixel,平均跟踪速率25.31 f/s。在双目视觉移动机器人平台上对随机运动目标进行跟踪实验,在背景干扰、遮挡、目标旋转和快速运动等组合情况下,跟踪成功率97.4%,跟踪距离平均相对误差为4.05%。

关键词: 目标跟踪;时空上下文;特征点跟踪;双目视觉;目标测量

中图分类号: TP391.4 TH72 **文献标识码:** A **国家标准学科分类代码:** 510.4050

Object tracking and location with spatio-temporal context assisted by key points

Zhai Jingmei, Liu Kun

(School of Mechanical and Automotive Engineering, South China University of Technology, Guangzhou 510640, China)

Abstract: Aimed at the problems of dynamic target tracking failure in the situation of fast motion and target occlusion, a spatio-temporal context tracking algorithm is proposed based on key points. Firstly, the key points of the target are extracted, and the predicted location of the object is obtained by combining key points matching with optical flow tracking. Then, the relationship model between key points change rate and spatio-temporal context model updating rate is established to control the update rate in real-time. In this way, the introduction of erroneous information can be prevented. Finally, a local context appearance model is constructed in the predicted location region, and the correlation between the spatio-temporal context model and the local context appearance model is computed to obtain the confidence map. Furthermore, the target is located accurately. The algorithm is validated in the test video, the highest average tracking success rate is 60% and the minimum average center error is 26.14 pixel. Compared to the 4 types of current major algorithms, the comprehensive performance of the proposed algorithm is superior to other methods, whose average tracking success rate is 90%, average center point error is 7.47 pixel and the average tracking rate is 25.31 frames per second. In the case of background interference, occlusion, target rotation and rapid motion, the mobile robot with binocular vision is used to track the random moving target. The success rate is 97.4%, and the average relative error of tracking distance is 4.05%.

Keywords: object tracking; spatio-temporal context; key points tracking; binocular vision; target measurement

0 引 言

目标在线跟踪在军事侦查、视频监控、行为识别、人机交互和移动机器人跟踪导航等领域有着广泛的应

用^[1-2]。跟踪算法中一个好的外观模型,可以使得跟踪在局部或全局遮挡、光照变化、姿态变化、尺度变化等表现出更好的鲁棒性,目标跟踪根据外观模型可分为产生式和判别式跟踪算法。产生式跟踪算法^[3-4]主要对目标的外观数据进行描述和建模,在后续帧中搜索与模型重构

误差最小的匹配区域作为目标位置,由于忽略了背景信息,在遇到遮挡等情况下模型错误更新容易产生目标漂移。判别式跟踪算法^[5-6]将跟踪问题转化为分类问题,不是对目标外观模型的刻画,而是充分利用背景与前景的信息将两者最大限度区分开,完成目标跟踪任务。

近年来基于相关滤波的跟踪算法不断被提出^[7],通过滤波模板与输入图像卷积,在响应值最大处得到目标位置,由于在频率域计算,该方法具有较好的实时性。基于该原理,文献[8]提出时空上下文跟踪算法(spatio-temporal context learning, STC),通过目标上下文环境关系,构建滤波模型,并在线更新模型,再根据与输入图像卷积的输出响应确定目标位置。这种基于目标与背景信息相关性的跟踪方法^[9],在目标部分遮挡和平面旋转时取得了较好的效果。但该方法以固定更新率更新模型(即每帧以固定的权值引入新模型),在目标长时间遮挡时易使模型错误更新产生目标漂移,由于其跟踪是基于帧间目标运动小的假设,在目标快速运动时会导致跟踪失败。在 STC 算法基础上,文献[10-11]针对跟踪精度对算法进行了改进,提高了跟踪器的性能,但没有解决目标遮挡和快速运动问题。针对目标遮挡问题, Sun S 等人^[12]采用高低更新率更新时空上下文模型,获得两幅对应的置信图,再搜索出最佳目标位置,但固定的高低更新率缺乏自适应性;张红颖等人^[13]在遮挡时使用级联分类器的辅助模型对时空上下文跟踪算法的有效性进行验证,防止错误更新,但该方法计算复杂耗时;刘威等人^[14]运用目标位置置信图峰值控制更新率,但目标快速运动模糊时也会引起置信度峰值变低,导致模型更新时可靠性较低。针对目标快速运动,文献[14]利用光流对随机标记的跟踪点进行运动估计,该方法在目标消失后重现时,目标点的光流跟踪失效,无法重新搜索目标恢复跟踪。

针对动态目标跟踪中出现的目标快速运动和遮挡问题,本文提出了一种利用特征点进行辅助定位和实时调控时空上下文模型更新的跟踪算法。引入特征点模型对目标局部特征进行描述,在目标大部分遮挡时仍能追踪到部分特征点。特征点数量的变化直接表征了目标遮挡的程度,由此根据特征点变化率自适应更新时空上下文模型。在目标快速运动、丢失重现时,根据特征点模型设计的检测机制可快速追踪到目标的大致区域,再结合时空上下文信息在该区域进一步精确定位目标,提高了定位精度。

1 特征点辅助的时空上下文跟踪算法

时空上下文跟踪算法^[8]在频域内计算进行目标跟踪,具有较好的实时性。但存在以下问题:1)当目标快速运动时,帧间目标位置变化较大。若仅根据前一帧目标位置在其附近区域搜索当前位置,导致帧间目标位置变

化较大,实际位置可能离开了预设的局部区域,导致无法在该区域内找到目标而跟踪失败。2)时空上下文模型以固定更新率进行更新,如果目标遮挡,此时的空间上下文模型会产生误差,持续以固定更新率将其引入时空上下文模型中,必然导致原时空上下文模型被错误更新。当目标长时间遮挡,这种错误更新的逐步积累最终会导致目标位置漂移。

针对问题1),提出了基于目标特征点的目标追踪,通过特征点目标位置预估,重新获取上下文模型的搜索区域。针对问题2),采用目标特征点个数的变化来表征目标受遮挡的程度,引入更新率调节参数,建立特征点变化率和时空上下文模型更新率关系,以此适应动态目标状态,实时修正更新率,确保跟踪可靠性。

1.1 基于特征点的目标位置预估

1.1.1 目标模型初始化

针对初始帧图像 I_1 ,进行二进制鲁棒特征点(binary robust invariant scalable keypoints, BRISK)检测,根据特征点在初始目标框内外,分离出前景和背景部分的特征点,同时生成前景和背景特征点的 BRISK 特征描述子^[15]。如图1所示,目标矩形框内的黑色点为前景点,矩形框外的白色点为背景点。通过对 I_1 帧图像的特征点提取和特征描述获得如下的目标模型:

$$O = \{(r_i, f_i, flag)\}_{i=1}^{N_o} \quad (1)$$

式中: r_i 是特征点 i 在初始帧的位置($r_i \in \mathbf{R}^2$), $f_i \in \{0, 1\}^d$ 为其对应的二进制特征描述子, d 表示特征维数, $flag$ 为标志位, 0、1 分别表示背景点与前景点, N_o 代表目标 O 特征点总数。考虑到特征点过多将增加耗时,较少又不足于表征目标,采用多个高低阈值进行特征点提取,选择最佳阈值使初始化目标特征点个数在 200 ~ 300。

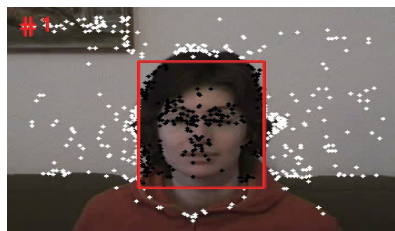


图1 目标模型初始化

Fig. 1 Initialization of the object model

1.1.2 基于光流和特征点匹配的目标追踪

在 $I_t (t > 1)$ 帧图像中,目标追踪关键工作在于寻找与目标模型 O 中对应的特征点 K_t 。

$$K_t = \{(a_i, m_i)\}_{i=1}^{N_k} \quad (2)$$

式中: a_i 表示 I_t 中第 i 个特征点的坐标位置, m_i 表示其在目标模板 O 中对应点的索引值,通过索引值可以确定每一个特征点在目标模型中的对应点。 K_t 在下文由类中值

流跟踪^[16]方法和特征点匹配获得。

为了适应目标的外观变化,采用光流对特征点进行连续追踪。在 I_t 帧中利用金字塔 Lucas-Kanada 光流法^[17]计算 I_{t-1} 帧中前景特征点 P_{t-1} 在 I_t 帧中的对应点 P_t ,再对 I_t 帧中对应点 P_t 计算逆向光流寻找其在 I_{t-1} 帧中的对应点 P_{t-1}^* 。在 P_t 中剔除逆向光流中未匹配的点和 $\|P_{t-1} - P_{t-1}^*\|$ 大于阈值的点,获得有效的跟踪点 T_t 。

利用 BRISK 特征点检测提取 I_t 帧中的特征点,并计算其 BRISK 特征描述子,对于获得的特征点 Q_t 。

$$Q_t = \{(a_i, f_i)\}_{i=1}^{N_{Q_t}} \quad (3)$$

再与 I_t 帧的目标模型 O 中的特征点描述子逐一进行比较,计算其汉明距离,公式如下:

$$D(f^1, f^t) = \sum_{j=1}^d \text{XOR}(f_j^1, f_j^t) \quad (4)$$

式中: f_j^1 表示 I_1 帧目标模型 O 中特征点的第 j 维特征值, f_j^t 表示 I_t 帧特征点 Q_t 的第 j 维特征值。 $D(f^1, f^t)$ 表示 I_1 帧目标模型 O 中特征向量与 I_t 帧 Q_t 特征向量的 Hamming 距离, XOR 为异或运算。

对于 Q_t 中的每个特征点运用式(4)计算其与 O 中特征描述子的距离,并返回测量距离小于 γ 的一个最优匹配点和次优匹配点,如果最优匹配点的测量距离非常小,而次优匹配点的测量距离相对较大,那么最优匹配点是安全可靠的,如果最优和次优匹配点的测量距离比值大于 η 时,表明两个候选匹配点的测量距离相近,选择其中之一作为匹配点很可能出错,剔除此不安全的匹配点。同时利用各特征点对应的索引值 m_i ,在 Q_t 中去除匹配到模板 O 中背景的点,得到有效的前景匹配点 M_t 。通过与类中值流跟踪法获得的特征点进行融合则获得 I_t 帧中目标特征点 $K_t = \{T_t \cup M_t\}$ 。

为了排除误匹配的孤立特征点的干扰,对获得的特征点 K_t 进行层次凝聚聚类^[18],类间距离小于 δ 的类进行合并,将获得的最大类作为最终的目标特征点 V ,根据特征点计算出目标中心位置 $L^* = \frac{1}{n} \sum_{i=1}^n V_i$,其中 $n = |V|$ 为 V 中特征点个数,FaceOcc1 序列中第 80 帧和 404 帧层次聚类前后特征点变化如图 2 所示,通过选择最大类,图 2 中书本上的非目标点被全部剔除,排除了孤立部分噪声点的干扰。图 3 所示为基于特征点跟踪的粗定位结果,由特征点求取的最小外接矩形可以看出,在第 80、404 帧发生部分遮挡和第 134、630 帧发生大面积遮挡情况下,该目标点跟踪方法均获得了较好的粗定位效果。

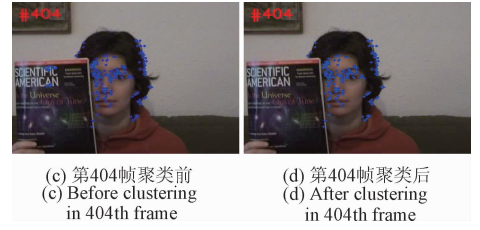
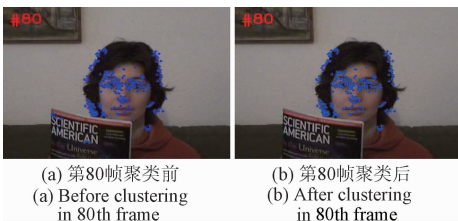


图 2 第 80 帧和 404 帧聚类前后特征点变化

Fig. 2 Changes of key points before and after clustering in 80th and 404th frames

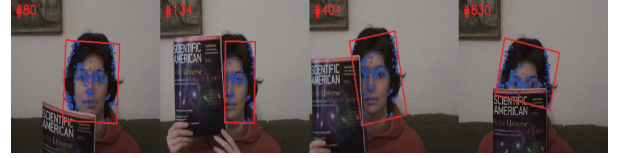


图 3 目标点跟踪结果

Fig. 3 Tracking results of the key points

1.2 时空上下文模型自适应更新

对于初始帧给定的目标位置 x 和目标位置中心 x^* , 其上下文特征集定义为 $X^c = \{v(z) = (I(z), z) \mid z \in \Omega_c(x^*)\}$, 其中 $I(z)$ 表示在 z 点的灰度值, $\Omega_c(x^*)$ 表示 x^* 为中心的局部区域。目标跟踪问题为计算估计目标位置 o 似然的置信图 $c(x)$:

$$c(x) = P(x \mid o) = \sum_{v(z) \in X^c} P(x, v(z) \mid o) = \sum_{v(z) \in X^c} P(x \mid v(z), o) P(v(z) \mid o) \quad (5)$$

同时置信图 $c(x)$ 的定义式为:

$$c(x) = b \cdot \exp\left(-\left|\frac{x - x^*}{\alpha}\right|^\beta\right) \quad (6)$$

式中: b 为归一化系数, α 是尺度参数, β 是形状参数,文献[8]中论证得出取 $\beta = 1$ 时能获得最佳效果。条件概率 $P(x \mid v(z), o)$ 表示对于给定的待选目标和其上下文特征,目标出现在点 x 的概率为:

$$P(x \mid v(z), o) = h^{sc}(x - z) \quad (7)$$

先验概率 $P(v(z) \mid o)$ 建模目标局部上下文外观模型,表示某一上下文特征属于目标的概率,用于选择与目标外观相似的上下文。对与目标中心位置不同距离的点赋予不同的权重,构建局部上下文外观模型为:

$$\begin{cases} P(v(z) \mid o) = I(z) \omega_\sigma(z - x^*) \\ \omega_\sigma(z - x^*) = a \cdot \exp\left(-\frac{(z - x^*)^2}{\sigma^2}\right) \end{cases} \quad (8)$$

式中: $\omega_\sigma(z - x^*)$ 表示权重函数, a 为归一化参数, σ 为外观模型尺度参数。根据式(5)~(8)可得以下关系式:

$$b \cdot \exp\left(-\left|\frac{x - x^*}{\alpha}\right|^\beta\right) = \sum_{v(z) \in X^c} h^{sc}(x - z) I(z) \omega_\sigma(z - x^*) = h^{sc}(x) \otimes (I(x) \omega_\sigma(x - x^*)) \quad (9)$$

式中: $\otimes$ 表示卷积运算, 为了加快计算速度, 将式(9)进行快速傅里叶变换 (fast Fourier transform, FFT), 在频率域内计算并得到空间上下文模型 $h^{sc}(x)$:

$$h^{sc}(x) = F^{-1} \left(\frac{F \left[b \cdot \exp \left(- \left| \frac{x - x^*}{\alpha} \right|^\beta \right) \right]}{F[I(x) \omega_\sigma(x - x^*)]} \right) \quad (10)$$

式中: $F(\cdot)$ 表示 FFT, $F^{-1}(\cdot)$ 表示快速傅里叶逆变换 (inverse fast Fourier transform, IFFT)。

为了适应跟踪环境的变化, 需要以某一更新率引入当前空间上下文模型, 完成时空上下文的学习更新。若采用固定更新率更新时空上下文模型, 当目标出现遮挡时将导致错误更新, 目标位置会出现漂移, 即使遮挡结束也无法重新锁定目标。利用 1.1 节中特征点定位的方法对 FaceOcc1 序列进行实验, 检测到的目标特征点数量变化如图 4 所示。

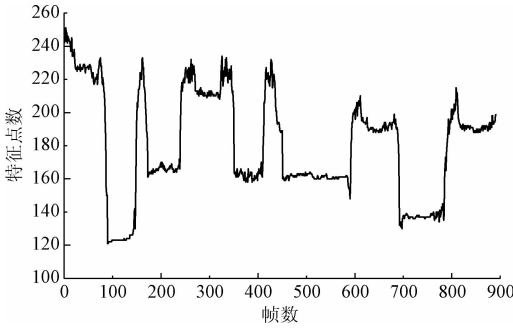


图 4 遮挡情况下特征点数量变化

Fig. 4 Changes of key points under occlusion

从曲线中可以看出, 在前 20 帧左右, 目标特征点个数维持在一个较高值, 第 20 帧后出现轻微遮挡后有一个急剧下降, 遮挡过程中维持一个较稳定的值。第 85 帧后出现大面积遮挡, 特征点个数迅速下降到一个较低值。第 150 帧后遮挡结束, 特征点个数又恢复到一个较高值。在之后的遮挡过程中特征点数量变化有着类似的规律, 遮挡时特征点个数下降, 遮挡越严重, 特征点个数下降的越多, 遮挡结束特征点个数又会恢复到较高值。

通过图 4 的实验表明特征点的数量变化可以表征目标受遮挡的程度, 遮挡越严重, 检测到的目标特征点数目就会越少。设计更新机制如式(11): 引入更新率调节参数 ω_t 实时修正更新率为 $\rho\omega_t$, 当特征点数目减少时, 降低更新率调节参数, 避免目标区域相关信息的错误引入。当特征点数目较多时, 增加更新率调节参数, 极大的增强了引入模型 h_t^{sc} 的可靠性, 增强了算法抗遮挡性能和跟踪的稳定性。

$$\begin{cases} H_{t+1}^{sc} = h_t^{sc} & t = 1 \\ H_{t+1}^{sc} = (1 - \rho\omega_t)H_t^{sc} + \rho\omega_t h_t^{sc} \\ \omega_t = \begin{cases} 0, & \lambda_t < 0.3 \\ \mu\lambda_t, & 0.7 > \lambda_t \geq 0.3 \\ \lambda_t, & \lambda_t \geq 0.7 \end{cases} & t > 1 \end{cases} \quad (11)$$

式中: H_{t+1}^{sc} 为第 $t+1$ 帧的时空上下文模型, ρ 为更新率, λ_t 表示第 t 帧中目标特征点个数与初始未遮挡时目标特征点个数的比值, μ 为微调参数。实验表明当比值小于 0.3 时遮挡很严重, 此时更新权重为 0, 模型不再更新; 当比值在 0.3 ~ 0.7 时为中度遮挡, 这种情况下 μ 取 0.375 对更新率进行微调; 比值大于 0.7 时为轻微遮挡。

1.3 基于时空上下文信息的目标位置获取

目标位置的精确定位采用求取置信图最大值位置, 以 1.1 节中特征点跟踪获得的中心位置 L^* 为中心, 在目标两倍大小的局部区域内对上下文外观特征进行建模:

$$P(v(z) | o) = I(z) \omega_\sigma(z - L^*) \quad (12)$$

对于第 $t+1$ 帧图像, 结合第 t 帧下由式(11)获得的时空上下文模型 H_{t+1}^{sc} , 求取当前帧的置信图, 以 $t+1$ 帧时的目标位置置信图中极大值的位置作为该帧目标位置, 目标精确位置为:

$$x_{t+1}^* = \underset{x \in \Omega_c(L_{t+1}^*)}{\operatorname{argmax}} c_{t+1}(x) =$$

$$\underset{x \in \Omega_c(L_{t+1}^*)}{\operatorname{argmax}} F^{-1} \{ F(H_{t+1}^{sc}) \otimes F[I_{t+1}(x) \omega_\sigma(x - L_{t+1}^*)] \} \quad (13)$$

式中: $\otimes$ 表示矩阵点乘运算, L_{t+1}^* 为 $t+1$ 帧基于特征点预估的目标中心位置, $\Omega_c(L_{t+1}^*)$ 表示以 L_{t+1}^* 为中心的局部区域。对于获得的最终目标位置 x_{t+1}^* , 利用式(10)求取 h_{t+1}^{sc} , 根据特征点变化率运用式(11)进行时空上下文模型学习更新得到 H_{t+2}^{sc} , 完成连续跟踪的过程。跟踪过程的同时为了适应目标距离变化, 结合置信图信息对目标尺度大小进行同步更新^[8]。

1.4 特征点辅助的时空上下文跟踪定位

本文跟踪算法框架如图 5 所示, 首先通过在初始帧建立目标特征点模型, 并获得初始帧目标点。在 $t+1$ 帧中提取 BRISK 特征点与初始目标模型匹配, 在全局内搜索目标, 基于该匹配可在目标遮挡和丢失后重新检测到目标。跟踪过程中目标相对初始目标模型会发生一定的变化, 为了适应目标外观变化, 采用前逆向光流法进行特征点跟踪。两种方法融合增强目标特征点的鲁棒性。再对特征点集进行层次聚类, 去除噪声点获得 $t+1$ 帧时的目标点。根据聚类后的点集, 对目标中心位置进行预估, 在预估位置局部区域求取局部上下文外观模型, 再与时空上下文模型 H_{t+1}^{sc} 进行卷积得到 $t+1$ 帧置信图, 在置信图最大值位置输出最终目标位置。由最终目标位置计算

$t+1$ 帧的时空上下文模型 h_{t+1}^{sc} , 根据层次聚类后特征点数相对目标初始特征点数的变化率得到 $t+1$ 帧时的更

新率 $\rho\omega_{t+1}$, 再计算更新后的时空上下文模型 H_{t+2}^{sc} , 下一帧循环以上过程, 进而完成目标的连续跟踪。

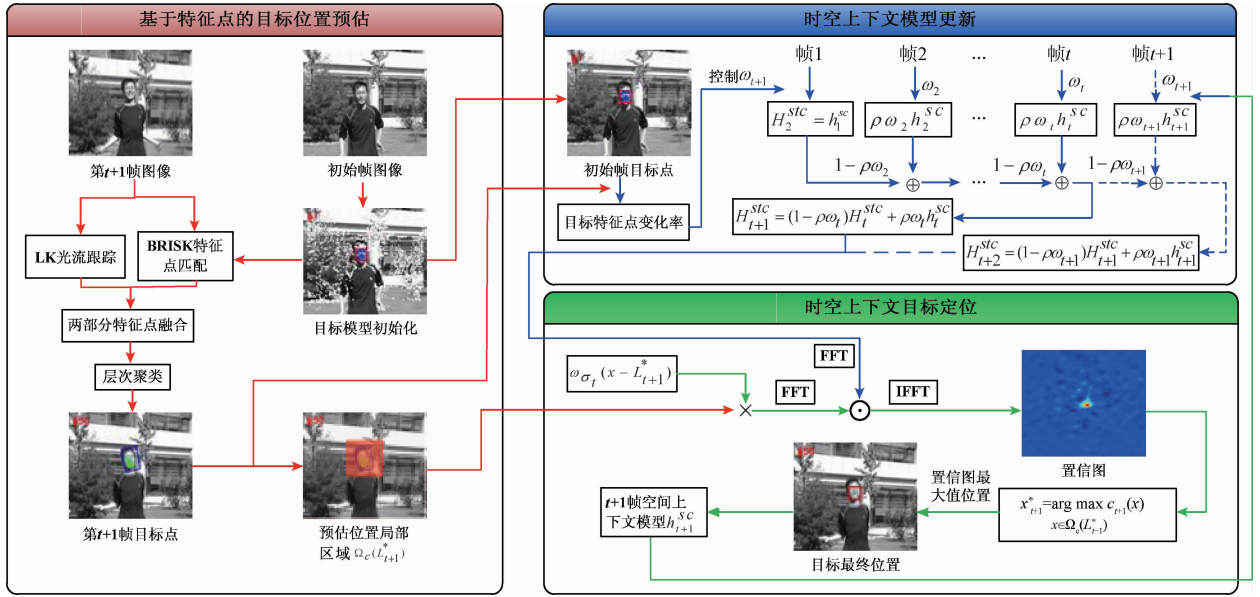


图 5 跟踪算法框架

Fig. 5 Framework of the tracking algorithm

2 算法性能测试实验

实验选用了视频测试集中 6 段具有不同挑战性的视频测试序列,除了和传统的 STC 算法^[8]、还选用了当前比较热门的快速压缩跟踪 (fast compressive tracking, CT) 算法^[19]、跟踪学习检测 (tracking-learning-detection, TLD) 算法^[20]、在线增强视觉 (on-line boosting and vision, OAB)^[21] 跟踪算法作比较,实验平台采用 Inter Core i5, CPU 主频 2.5 GHz 的 PC,利用 OpenCV2.4.9 在 VS2010 上进行调试。实验中学习更新率 ρ 取 0.075,微调参数 μ 取 0.375, γ 取 0.25, η 取 0.8, δ 取 20。

选用跟踪成功率和中心位置误差作为算法性能评估标准,跟踪成功率定义为跟踪成功帧数占总帧数的比值,对于跟踪分值大于 0.5 的认为跟踪成功。跟踪分值 S_{score} 和中心位置误差 e 计算如下:

$$\begin{cases} S_{score} = \frac{R_T \cap R_C}{R_T \cup R_C} \\ e = \sqrt{(x_T - x_C)^2 + (y_T - y_C)^2} \end{cases} \quad (14)$$

式中: R_T 和 R_C 分别表示跟踪结果区域和真实结果区域。 (x_T, y_T) 表示跟踪结果中心坐标, (x_C, y_C) 表示真实目标中心坐标。6 个测试序列跟踪效果如图 6 所示,分别展现了背景干扰、全局和部分遮挡、快速运动与运动模糊、目标旋转等情况下的算法跟踪效果。跟踪成功率和跟踪

平均中心位置误差如表 1 所示,可以看出本文改进后的算法在平均跟踪成功率上比 STC 算法有了显著的提高,在对比跟踪算法中效果最优,平均跟踪成功率达到了 90%。表 2 所示为各算法在当前测试平台下的跟踪速率,STC 算法跟踪速率最快,但其对快速运动目标跟踪 (Jumping) 的成功率只有 5%,远低于本文算法且跟踪误差也非常大;CT 算法跟踪速率虽高于本文算法,但在背景干扰 (CarDark) 和快速运动 (Jumping) 下跟踪成功率分别为 0 和 1%,大面积遮挡情况下 (Coke) 跟踪成功率为 9%,平均跟踪成功率仅有 32%,远低于本文算法;本文算法平均跟踪速率为 25.31 f/s,在较高的跟踪成功率和跟踪精度下,满足了实时性要求。图 7 所示为不同算法相应的跟踪中心位置误差曲线,本文算法在各类测试视频中都具有较小的位置误差,平均中心位置误差为 7.47 pixel。

结合各视频的特性对图 7 进行分析如下。

1) 目标旋转变化:在图 6(a) 中第 1 061 帧玩具转过较大角度,时空上下文关系急剧下降,STC 算法出现目标漂移,并随着模板持续更新在第 1 108 帧时已经完全无法恢复对目标的跟踪。本文算法通过特征点进行初步定位和控制模型更新,成功率最高。

2) 背景干扰:图 6(c) 中 CT 算法在第 260 帧之前早已跟踪失败且持续无法恢复,TLD 算法在第 260 帧时由于目标周围区域出现了相似车辆的干扰导致跟踪出现偏差,并在第 284,315 帧时完全跟错,而 OAB、STC 和本文算法跟踪效果都很理想。



图 6 不同算法跟踪效果
Fig. 6 Tracking results of different methods

表 1 跟踪成功率 η 与平均中心误差 $\bar{e}$
Table 1 Tracking success rate η and average center location error $\bar{e}$

图像序列	CT($\eta/\bar{e}$)	TLD($\eta/\bar{e}$)	STC($\eta/\bar{e}$)	OAB($\eta/\bar{e}$)	本文算法($\eta/\bar{e}$)
Sylvester	0.83/8.7	<u>0.93/7.3</u>	0.65/27	0.68/15	0.95/6.94
Woman	0.16/114	0.17/139	0.92/12.5	0.61/32	<u>0.90/5.3</u>
CarDark	0.00/119	0.53/27	1.00/4.6	0.95/3	<u>0.99/2.85</u>
FaceOccl	0.85/25.6	0.83/27	0.24/70.7	<u>0.91/25</u>	0.99/12.13
Jumping	0.01/47.9	<u>0.85/6</u>	0.05/52	0.05/46	0.98/3.43
Coke	0.09/40.5	<u>0.29/24.9</u>	0.28/17.7	0.17/35.86	0.56/14.17
平均值	0.32/59.3	<u>0.60/38.5</u>	0.52/30.8	0.56/26.14	0.90/7.47

注:加粗字体为该行最优,划横线字体为该行次优

表 2 各算法跟踪速率
Table 2 Tracking speed of each algorithm

图像序列	CT	TLD	STC	OAB	本文算法
Sylvester	37.11	18.35	70.2	4.13	26.15
Woman	48.5	23.97	77.61	8.21	25.34
CarDark	53	21.7	85.3	14.8	31.6
FaceOccl	31.13	17.85	63.7	2.89	19.03
Jumping	43.80	15	70.8	7.32	30.4
Coke	42.03	17.4	77.4	4.08	19.31
平均值	42.6	19.05	74.17	6.91	25.31

(f/s)

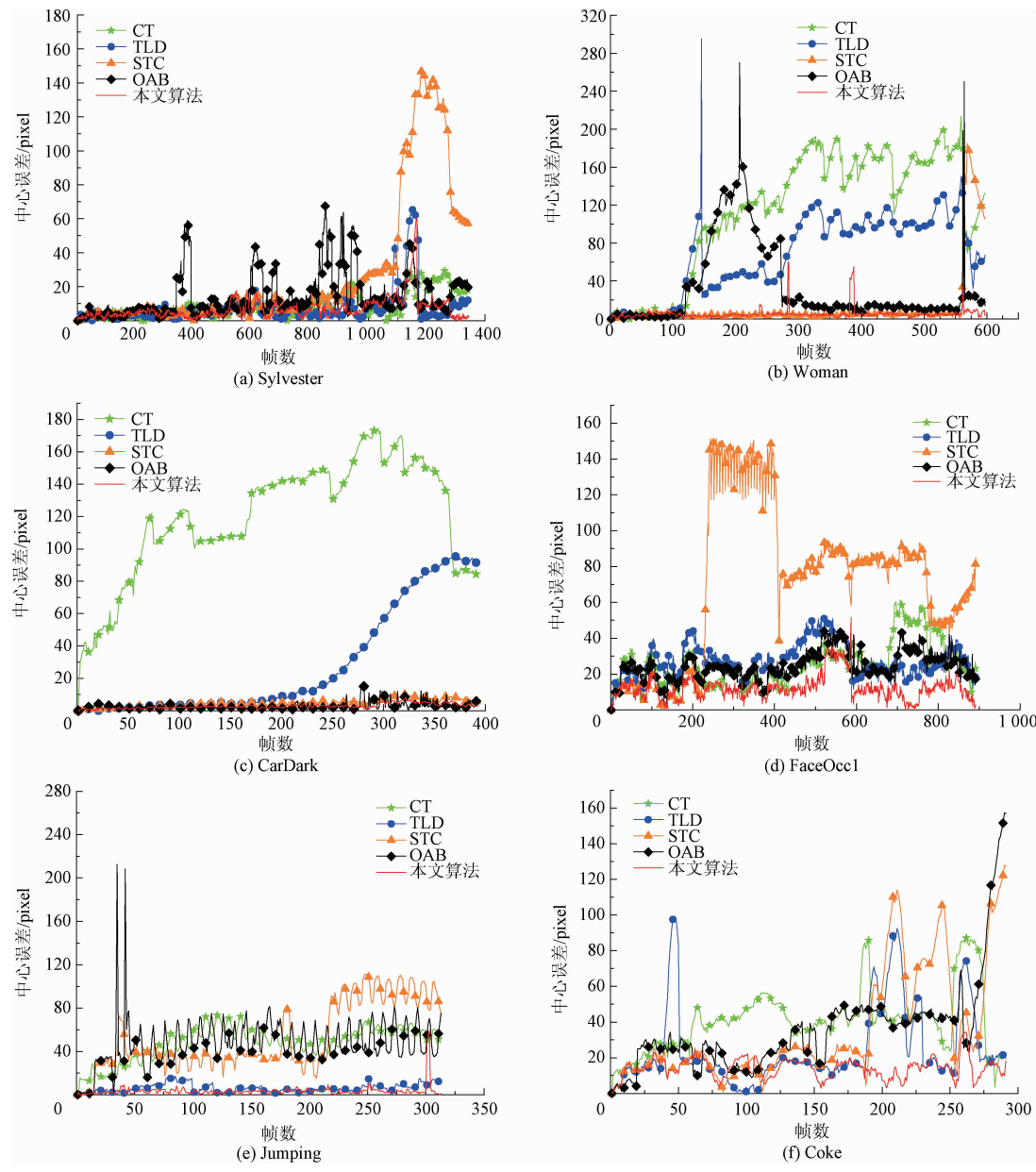


图 7 视频序列跟踪中心偏差
Fig. 7 Tracking error of video sequence

3) 局部遮挡:在图 6(b)中短时局部的遮挡 STC 算法仍取得较好的跟踪效果,在第 127 帧出现部分遮挡时 TLD、OAB 和 CT 算法都失效了,之后 OAB 算法恢复,但 CT 和 TLD 算法没有恢复跟踪。在第 563 帧时镜头突然拉近,目标偏移上一帧位置较大,STC 算法跟踪失败,本文算法还是准确跟踪到目标。图 6(d)人脸出现长时间大面积遮挡,第 232 帧后由于长时间遮挡,STC 模型更新错误导致跟踪失败。第 562 帧出现大面积持续遮挡, OAB 算法出现了偏移, TLD 算法也只检测到剩余未遮挡部分,跟踪效果较差。第 736 帧时 CT 算法出现位置偏移,本文算法在整个遮挡过程中准确的跟踪到了目标。

4) 完全遮挡:图 6(f) 存在目标旋转和多次完全遮挡,在第 205 帧经过完全遮挡和旋转变换后,STC 算法完全丢失目标无法恢复跟踪, TLD 也丢失了目标但在后续帧中恢复了跟踪, OAB、CT 算法在目标旋转时产生了位置漂移。第 258 帧目标完全被绿叶遮挡,在第 290 帧目标重新出现时 STC、OAB 算法跟踪完全失效, CT 和 TLD 算法跟踪到目标但出现了较大偏差,目标重新出现时本文算法依然准确跟踪到目标。

5) 快速运动与运动模糊:图 6(e) 中第 16 帧中人上下跳动时目标位置快速变化,人脸部特征变的模糊, STC、OAB、CT 算法都跟踪失败,之后无法恢复对目标的

跟踪。第 95、178 帧时,TLD 由于目标模糊程度较大,跟踪框出现了较大偏移,本文算法在根据特征点粗定位的基础上利用时空上下文的关系更精确的定位了目标,取得了比 TLD 更好的跟踪效果。

各项指标测试中,相比当前主流算法,本文算法保持最优和次优的成绩,综合测试指标中跟踪效果最佳,在局部遮挡、背景干扰、快速运动和目标旋转跟踪中,均达到 90% 以上的成功率。对于全局遮挡情况下,跟踪成功率相比其他算法大大提高,成功率为 56% 是由于目标出现长时间全局遮挡,算法无法进行位置预估,但当目标一旦重新出现时,算法又能及时跟踪和精确定位目标。

3 移动机器人跟踪实验

针对移动机器人目标跟踪过程中,需要对目标距离实时测量,实验采用双目测距系统。首先利用张正友标定法标定左右相机,获得相机内外参数及畸变参数,根据获得的参数完成立体标定,计算出右相机相对左相机的旋转平移参数。利用标定的参数对输入的图像对进行立体校正,获得行对齐的双目校正图像对。再运用基于模板匹配的视差计算方法,将左相机下跟踪获得的目标区域作为匹配模板,在校正后的右相机图像中匹配,进行匹配的同时引入视差范围约束减小搜索区域,使其在同行的局部区域内搜索,从而提高匹配精度和速度,快速获得目标区域中心的视差。对于图像中目标中心点 $g(u, v)$,利用双目测距原理恢复目标中心点的三维坐标点 $G(x, y, z)^{[22]}$ 。

$$\begin{cases} x = (u - c_x) / w \\ y = (v - c_y) / w \\ z = f / w \\ w = -d / T_x \end{cases} \quad (15)$$

式中: d 表示视差, f 表示焦距, c_x 和 c_y 分别表示主点(主光线与像平面的交点)在左相机图像上的 x 和 y 坐标, T_x 为两相机的水平位移。

移动机器人动态目标跟踪工作原理如图 8 所示,机器人采用麦克纳姆轮,通过 4 个轮子的转速转向适当组合,实现机器人在平面上三自由度的全方位移动。双目视觉系统置于机器人前端,采用平行方式安装,基线距离约 5 cm,通过双目系统采集的图像传到 PC 机进行跟踪测量处理。根据设定的距离预期值,机器人控制系统运用视觉系统提供的目标坐标信息形成控制信号,并发送给 STM32 控制器,进而控制驱动器驱动电机完成相应的跟踪动作,使目标和移动机器人保持在一定的距离内,完成目标跟踪任务。

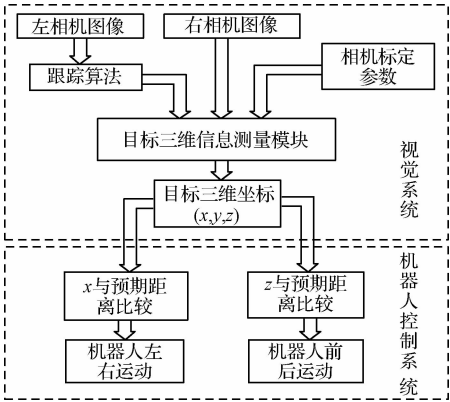


图 8 移动机器人系统框架

Fig8 Framework of the mobile robot system

实验中跟踪目标选用扫地机器人,运动过程中若与其他物体碰撞会做旋转运动,路径随机具有不确定性,为了在实验中体现其抗干扰性能,在扫地机器人顶部贴上二维码进行跟踪,并在环境中设置类似二维码进行背景干扰。实验分别对光线变化、旋转、干扰、快速运动和遮挡等综合情景进行跟踪测试,跟踪效果如图 9 和 10 所示,图片左上角的数字为帧数,右上角的字母 L 表示左相机图像,目标跟踪结果由矩形框表示。



图 9 环境亮度变化

Fig. 9 Brightness change of environment



(a) 旋转和背景干扰
(a) Rotate and background disturb



(b) 旋转、背景干扰和快速运动
(b) Rotate, background disturb and fast moving



(c) 旋转、背景干扰和遮挡
(c) Rotate, background disturb and occlusion

图 10 复杂环境下跟踪

Fig. 10 Tracking in complex environment

图 9 中扫地机器人从光线较强的区域逐渐走向光线较暗的区域,整个过程中,移动机器人准确跟踪。图 10 场景中设置了与扫地机器人相似的二维码,结果表明算法在实际运用中具有较强的抗干扰能力。图 10(a)中在第 36、45 帧中,扫地机器人与书本进行碰撞,发生旋转,移动机器人依然跟踪准确;图 10(b)中在第 292 帧,在干扰背景下人为踹动扫地机器人,使其快速运动,目标出现拖影,仍成功定位到目标位置;图 10(c)中第 538 帧目标被逐渐遮挡,第 544 帧中被完全遮挡时目标丢失,548 帧时目标重新出现,移动机器人再次检测到目标,恢复对目标的跟踪。

对于图 10 复杂环境下的跟踪实验,实验共 681 帧,跟踪成功率 97.4%。在第 424 ~ 430 帧、540 ~ 547 帧、616 ~ 618 帧跟踪失败,共 18 帧,完全是由目标完全遮挡造成。实验跟踪预期距离为 130 cm,实际跟踪平均距离为 135.27 cm,平均相对误差 4.05%。

4 结 论

本文将目标特征点引入到传统时空上下文跟踪算法中,基于目标特征点及其变化,提出了一种有效的动态目标跟踪新算法。算法建立目标特征点模型,采用特征点匹配和光流跟踪方法在全局范围内搜索目标,两种方法融合增强匹配目标特征点的鲁棒性;对特征点集进行层次聚类,对目标中心先进行位置预估,有效解决了目标快速运动导致跟踪失败的问题;根据目标特征点数目变化,自适应地调节时空上下文模型的更新率,在目标遮挡情况下获取了更可靠的跟踪效果。算法在视频测试集上分别测试了背景干扰、遮挡、目标旋转和快速运动等 6 种情境下的跟踪效果:平均跟踪速度 25.31 f/s,具有较好的实时性;跟踪平均成功率达到 90%,平均跟踪误差为 7.47 pixel,相比其他主流算法,跟踪准确率最高、误差最小。算法在双目移动机器人平台上进行动态目标跟踪实验,在亮度变化、背景干扰、目标旋转、遮挡和运动模糊等综合情境环境中达到了理想的跟踪效果:跟踪成功率 97.4%、跟踪距离平均相对误差为 4.05%。

参考文献

- [1] 厉艳琨,毛建旭,刘仁明. 基于特征点的 3D 人脸姿态跟踪[J]. 电子测量与仪器学报, 2016, 30(4): 605-612.
- LI Y K, MAO J X, LIU R M. 3 D face pose tracking based on feature matching[J]. Journal of Electronic Measurement and Instrumentation, 2016, 30(4): 605-612.
- [2] 李朕阳,郎朗,陈孟元. 基于 SR-CKF 的移动机器人动态目标跟踪算法[J]. 电子测量与仪器学报, 2016,

30(8):1198-1205.

- LI ZH Y, LANG L, CHEN M Y. Dynamic target tracking algorithm for mobile robot based on square-root cubature Kalman filter[J]. Journal of Electronic Measurement and Instrumentation, 2016, 30(8): 1198-1205.
- [3] ZHANG T, LIU S, XU C, et al. Structural sparse tracking[C]. Computer Vision and Pattern Recognition, 2015:150-158.
- [4] 孔繁镔,王丹丹,沈秋. L1-L2 范数联合约束的鲁棒目标跟踪[J]. 仪器仪表学报, 2016, 37(3): 690-697.
- KONG F Q, WANG D D, SHEN Q. Robust object tracking via L1- L2 norm simultaneous constraint[J]. Chinese Journal of Scientific Instrument, 2016, 37(3): 690-697.
- [5] NAM H, HAN B. Learning multi-domain convolutional neural networks for visual tracking[C]. Computer Vision and Pattern Recognition, 2016:4293-4302.
- [6] 苑晶,李阳,董星亮,等. 基于运动模式在线分类的移动机器人目标跟踪[J]. 仪器仪表学报, 2017, 38(3):568-577.
- YUAN J, LI Y, DONG X L, et al. Target tracking with a mobile robot based on on-line classification for motion patterns[J]. Chinese Journal of Scientific Instrument, 2017, 38(3):568-577.
- [7] DANELLJAN M, HAGER G, KHAN F S, et al. Learning spatially regularized correlation filters for visual tracking[C]. Proceedings of the IEEE International Conference on Computer Vision, 2015: 4310-4318.
- [8] ZHANG K, ZHANG L, LIU Q, et al. Fast visual tracking via dense spatio-temporal context learning[C]. Computer Vision-ECCV, 2014:127-141.
- [9] HENRIQUES J F, CASERIO 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.
- [10] WEI X G, ZHANG S, CHAN S C. A novel visual object tracking algorithm using multiple spatial context models and Bayesian Kalman filter[C]. IEEE International Symposium on Circuits and Systems, 2015:1034-1037.
- [11] 徐建强,陆耀. 一种基于加权时空上下文的鲁棒视觉跟踪算法[J]. 自动化学报, 2015, 41(11): 1901-1912.
- XU J Q, LU Y. Robust visual tracking via weighted spatio-temporal context learning[J]. Acta Automatica Sinica, 2015, 41(11):1901-1912.
- [12] SUN S, ZHANG H, YUAN D. Robust visual tracking

- with dual spatio-temporal context trackers[C]. Seventh International Conference on Graphic and Image Processing, 2015:1-5.
- [13] 张红颖, 郑轩. 基于双目标模型的时空上下文跟踪算法[J]. 光学精密工程, 2016, 24(5):1215-1223.
ZHANG H Y, ZHENG X. Spatio-temporal context tracking algorithm based on dual-object model [J]. Optics and Precision Engineering, 2016, 24 (5): 1215-1223.
- [14] 刘威, 赵文杰, 李成. 时空上下文学习长时目标跟踪[J]. 光学学报, 2016, 36(1):179-186.
LIU W, ZHAO W J, LI CH. Long-term visual tracking based on spatio-temporal context [J]. Acta Optica Sinica, 2016, 36(1):179-186.
- [15] NEBEHAY G, PFLUGFELDER R. Clustering of static-adaptive correspondences for deformable object tracking[C]. Computer Vision and Pattern Recognition, 2015:2784-2791.
- [16] KALAL Z, MIKOLAJCZYK K, MATAS J. Forward-backward error: Automatic detection of tracking failures[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2010:2756-2759.
- [17] LUCEY S, NAVARATHNA R, ASHRAF A B, et al. Fourier Lucas-Kanade algorithm[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(6):1383-1396.
- [18] XU D, TIAN Y. A comprehensive survey of clustering algorithms[J]. Annals of Data Science, 2015, 2(2): 165-193.
- [19] ZHANG K, ZHANG L, YANG M H. Fast compressive tracking[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014, 36(10):2002-2015.
- [20] KALAL Z, MIKOLAJCZYK K, MATAS J. Tracking-learning-detection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34 (7): 1409-1422.
- [21] GRABNER H, GRABNER M, BISCHOF H. Real-time tracking via on-line boosting[C]. British Machine Vision Conference, 2006:47-56.
- [22] 王浩, 许志闻, 谢坤, 等. 基于 OpenCV 的双目测距系统[J]. 吉林大学学报:信息科学版, 2014, 32(2): 188-194.
WANG H, XU ZH W, XIE K. et al. Binocular measuring system based on OpenCV[J]. Journal of Jilin University : Information Science Edition, 2014, 32 (2): 188-194.

作者简介



E-mail: mejmzhai@scut.edu.cn

Zhai Jingmei (Corresponding author) received her B.Sc. degree from Xi'an Jiaotong University in 1990, received her M.Sc. and Ph.D. degrees both from South China University of Technology in 1998 and 2003, respectively. Now she is a professor in South China University of Technology. Her main research interests include mechanical system modeling and optimization, information processing of electromechanical equipment, and artificial intelligence.



刘坤, 2015 年于南昌大学获得学士学位, 现为华南理工大学研究生, 主要研究方向为机器视觉及图像处理。

E-mail: kunliusg@163.com

Liu Kun received his B.Sc. degree from Nanchang University in 2015. Now he is a M.Sc. candidate in South China University of Technology. His main research interests include machine vision and image processing.