

运动模型引导的自适应核相关目标跟踪方法

陈 丹, 姚伯羽

(西安理工大学自动化与信息工程学院, 陕西西安 710048)

摘 要: 针对小型移动机器人对人体目标快速运动或遮挡导致的跟踪准确率降低甚至跟踪失败问题, 通过建立足部运动模型预测双脚位置信息, 获得核相关滤波(KCF, Kernel Correlation Filter)目标检测区域, 再结合输出响应峰值邻域相关检测, 提出了运动模型引导的自适应核相关滤波算法. 对实际拍摄的七组不同情况下的视频进行了足部目标跟踪实验, 结果表明运动模型引导的自适应响应 KCF 算法平均跟踪准确率最高, 且在短时间遮挡情况下的算法跟踪准确率也达到 86%, 明显高于自适应响应 KCF、BACF (Background Aware Correlation Filters) 以及 SAMF (Scale Adaptive kernel correlation filters with Multiple Features) 三种跟踪算法. 最后在 ROS (Robot Operating System) 下将所提算法应用于 Turtlebot 机器人目标跟踪测试, 成功克服了遮挡情况对足部跟踪带来的影响, 验证了所提算法具有较强的鲁棒性和实时性.

关键词: 目标跟踪; 运动模型; 核相关滤波; 自适应输出响应; 遮挡检测; 机器人操作系统

中图分类号: TP391 **文献标识码:** A **文章编号:** 0372-2112 (2021)03-0550-09

电子学报 URL: <http://www.ejournal.org.cn> **DOI:** 10.12263/DZXB.20200433

Adaptive Response Kernel Correlation Target Tracking Method Guided by Motion Model

CHEN Dan, YAO Bo-yu

(School of Automation and Information Engineering, Xi'an University of Technology, Xi'an, Shaanxi 710048, China)

Abstract: Aiming at the problem of low tracking accuracy and even tracking failure caused by fast motion or occlusion of human targets by small mobile robots, a foot motion model was established to predict the position information of feet, and the target detection region of kernel correlation filter (KCF) was obtained. A motion model guided adaptive kernel correlation filtering algorithm is proposed by combining with the output response peak neighborhood correlation detection. Foot tracking experiments were carried out on seven groups of videos under different scenarios. The results show that the average tracking accuracy of the adaptive response KCF algorithm guided by the motion model is the highest, and the tracking precision rate of the algorithm reaches 86% in the case of short-term occlusion, which is significantly higher than that of the adaptive response KCF, BACF and SAMF algorithm. Finally, the proposed algorithm is applied to the target tracking test of a Turtlebot robot under the ROS (Robot Operating System), which successfully overcomes the influence of occlusion on feet tracking, and verifies that the proposed algorithm has strong robustness and real-time performance.

Key words: target tracking; motion model; kernel correlation filter; adaptive response; occlusion detection; ROS

1 引言

随着人工智能等先进技术的加速发展, 机器人迅速从工业领域向服务行业渗透. 人体跟随机器人属于服务类机器人的一种, 可以被应用于机场、仓库、医院等场合的货物运输. 具有强鲁棒性、强实时性的跟踪算法是视觉跟随机器人完成跟随任务的先决条件. 相关滤

波类跟踪算法因为具有较高的准确率和优异的跟踪速度成为计算机视觉领域的研究热点之一. 2010 年 Bolme 等人提出的最小均方误差算法 (MOSSE)^[1] 最早将相关滤波引入到目标跟踪中. 2015 年 J. F. Henriques 等人提出了核相关滤波 KCF 跟踪算法, 在测试集 OTB (Object Tracking Benchmark) 下的跟踪准确率达到 0.74^[2]. Yi Wu 等人认为运动模型能使目标跟踪算法的性能得到

提升^[3]. 近年来结合运动模型的跟踪算法也不断被提出, K. Ratnayake 等人提出了一种基于动态运动模型和支持向量机(SVM)的任意目标跟踪方法^[4]. Wang Xiong-peng 等人提出了一种基于多运动模型的自校正跟踪方法, 解决视觉跟踪中的目标速度和方向突变等不规则运动问题^[5]. Wang Ying 等人提出了一种用于自主式水下机器人(AUV)的目标跟踪算法, 建立目标的典型运动模型和非线性随机运动模型^[6], 可以估计出声纳的运动轨迹. 陈婧等人在红外制导领域提出融合卡尔曼滤波和运动模型的改进核相关目标跟踪算法^[7]. 廖家才等人提出了基于自动驾驶的运动模型改进核相关滤波跟踪算法^[8], 使用卷积神经网络 YOLOv2 建立了多目标的检测跟踪容器, 融合了核相关滤波和目标运动模型对目标轨迹进行跟踪.

对于小型人体跟随机器人, 相机高度越高目标人可以越完整的出现在相机视野内. 小型移动机器人的相机都处于较低的高度且高度不可调整, 如果增大摄像头相对于水平线的仰角, 那么此时摄像头将存在很大的视野盲区^[9], 因此采用传统的上半身跟踪适用性较低, 而将人体足部作为跟踪目标更为合适, 因为摄像头可以安装在一个较低的位置, 能够在视野更大的范围内通过足部跟踪人体目标. 本文首先建立了人体足部运动模型, 基于运动模型提出自适应响应核相关滤波算法, 将响应图峰值邻域相关性作为遮挡判断依据, 解决了遮挡情况下的目标跟踪问题.

2 人体足部运动模型建立

人体双脚运动具有一定的周期性, 双腿带动双脚的摆动交替进行. 假设双脚均为刚体, 不难发现两脚各自中心点的在垂直方向的位移变化接近于低振幅的余弦波^[10]. 本文将人体行走时的足部运动分为两部分: 局部运动和全局运动. 其中两脚之间的相对运动为局部运动, 两脚整体运动为全局运动.

2.1 局部运动模型

本文拍摄了三组不同步速的人体行走视频. 选取从后方观测双脚各自的最低点位置计算两脚间的垂直位移, 统计三个视频中连续 100 帧图像中的两脚间垂直位移后得到图 1 所示曲线, 其中图 1(a)为慢速行走, 图 1(b)为中等步速行走, 图 1(c)为快速行走. 从图 1 可以看出, 人体在行走的过程中两脚间的实际垂直位移的变化可采用余弦函数拟合.

在某一帧图像中, 取左脚位置坐标为 (x_l, y_l) , 右脚位置为 (x_r, y_r) , 则两只脚在垂直方向上的相对位移为 $t_{zero}^k = t - 1 + \frac{|d^{t-1}|}{|d^t| + |d^{t-1}|} d^t \cdot d^{t-1} \leq 0$, 其中 t 表示帧数. 采用幅度和相位调制余弦函数对两只脚之间的垂直位

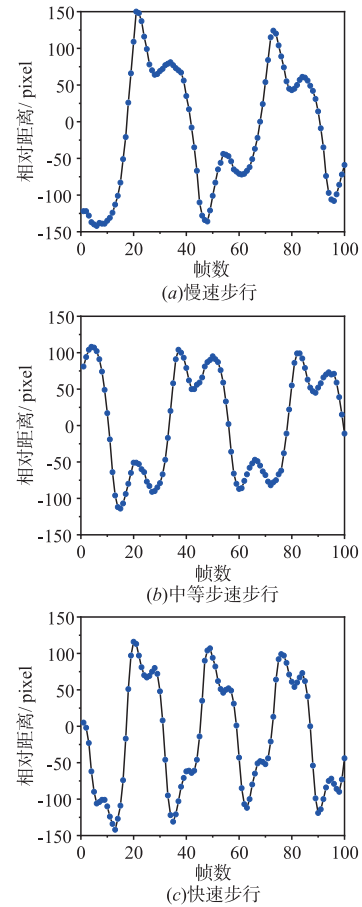


图1 行走过程中两个脚后跟之间的垂直距离

移进行建模, 计算公式如下:

$$d^t = A^t \cos\left(\frac{2\pi}{T^t} t + \varphi^t\right) \quad (1)$$

其中: A^t 为余弦函数振幅, T^t 为周期, φ^t 为相位.

在式(1)中, 周期可由 d^t 数据的两个连续过零点时间获得, 则过零点的时间为:

$$t_{zero}^k = t - 1 + \frac{|d^{t-1}|}{|d^t| + |d^{t-1}|} d^t \cdot d^{t-1} \leq 0 \quad (2)$$

其中, t_{zero}^k 表示第 k 个过零点时间, 周期 $T^t = 2 \times (t_{zero}^k - t_{zero}^{k-1})$ 则相位为:

$$\varphi^t = \begin{cases} \frac{\pi}{2} - \frac{2\pi}{T^t} t, & d^{t-1} > 0 \text{ 且 } d^t < 0 \\ \frac{3\pi}{2} - \frac{2\pi}{T^t} t, & d^{t-1} \leq 0 \text{ 且 } d^t \geq 0 \end{cases} \quad (3)$$

在研究了拍摄的多组视频后, 发现人体正常行走过程中双脚间水平的位移变化很小, 因此, 当前帧中的水平位移估计为 $l_{pred}^t = l^{t-1}$.

2.2 局部运动模型

采用两脚之间中点位置的变化建立人体足部全局运动模型. 人体行走状态在连续帧之间变化不快, 所以可假设连续两帧之间行走速度不变, 用表示第 t 帧两脚

之间连线的中点,如图 2 所示.

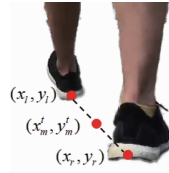


图2 双脚及中点坐标图

第 t 帧的中点预测位置可由前面两帧中点位置获得:

$$\begin{pmatrix} x_{mpred}^t \\ y_{mpred}^t \end{pmatrix} = 2 \cdot \begin{pmatrix} x_m^{t-1} \\ y_m^{t-1} \end{pmatrix} - \begin{pmatrix} x_m^{t-2} \\ y_m^{t-2} \end{pmatrix}, t \geq 2 \quad (4)$$

其中,从目标人后方观测的左脚和右脚中点预测位置分别表示为:

$$\begin{pmatrix} x_{lpred}^t \\ y_{lpred}^t \end{pmatrix} = \begin{pmatrix} x_{mpred}^t - \frac{l^t}{2} \\ y_{mpred}^t + \frac{d^t}{2} \end{pmatrix}, \begin{pmatrix} x_{rpred}^t \\ y_{rpred}^t \end{pmatrix} = \begin{pmatrix} x_{mpred}^t + \frac{l^t}{2} \\ y_{mpred}^t - \frac{d^t}{2} \end{pmatrix} \quad (5)$$

3 自适应响应核相关滤波算法

核相关滤波算法主要存在问题^[11,12]: (1) 检测结果不准确:在遮挡、形变、快速运动等情况下,目标响应会发生一定的变化,发生此类情况的帧数较多时将导致目标响应被污染,这种误差会传递到下一帧的分类器系数计算,形成误差积累,跟踪框将存在漂移风险;(2) 循环移位不准确:循环移位产生的样本只是目标运动的近似值,在面对遮挡和快速运动等情况时会变得不再可靠,循环移位带来的边界效应会对分类器的判断能力造成影响.

在 KCF 算法中,分类器是通过目标周围大小为 $m \times n$ 的图像块 z 训练获得. 假设图像块 z 的目标函数为 $f(z) = w^T z$,则分类器是通过最小正则化风险泛函进行训练:

$$\min_w \sum_i (f(x_i) - y_i)^2 + \lambda \|w\|^2 \quad (6)$$

其中 y_i 是回归标签, λ 是正则化参数防止过拟合, w 是权重. 将式(6)写成矩阵形式为:

$$\min_w \|Xw - y\|^2 + \lambda \|w\|^2 \quad (7)$$

其中 y 是期望输出的矩阵,也即回归标签矩阵, X 表示循环样本矩阵, X 的每一行表示一个样本 x_i .

对回归标签 y 进行自适应改变,在跟踪过程中不断使用目标响应 y_0 对回归标签修正,使二者的误差尽可能小. 因此,在式(7)后添加一个损失函数项:

$$\min_{w,y} \|Xw - y\|^2 + \lambda_1 \|w\|^2 + \lambda_2 \|y - y_0\|^2 \quad (8)$$

式中 λ_2 为正则化参数, $y = y_0 + n$, n 为噪声,且 $y \sim N(y_0, \text{diag}^{-1}(1/2\lambda_2))$.

令 $z^T = [w^T \ y^T]$, 则式(8)可写为:

$$f(z) = \| [X \ -I]z \|^2 + \lambda_1 \| [I \ 0]z \|^2 + \lambda_2 \| [0 \ I]z - y_0 \|^2 \quad (9)$$

式中, w 、 y 、 y_0 均为 n 维列向量, z 为 $2n$ 维列向量. 将二元线性回归求解转化为了二元线性回归求解,降低了计算复杂度. 采用梯度下降法对式(9)求解得:

$$\nabla_z f(z) = \begin{bmatrix} X^T X & -X^T \\ -X & I \end{bmatrix} z + \lambda_1 \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} z + \lambda_1 \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix} z - \lambda_2 \begin{bmatrix} 0 \\ I \end{bmatrix} y_0 \quad (10)$$

令式(10)为零,并将 $z^T = [w^T \ y^T]$ 代入,利用循环矩阵的傅里叶对角化性质可得:

$$\begin{bmatrix} \hat{w}^* \\ \hat{y}^* \end{bmatrix} = \lambda_2 \begin{bmatrix} \text{diag}(\hat{x} \odot \hat{x}^* + \lambda_1) & -\text{diag}(\hat{x}^*) \\ -\text{diag}(\hat{x}) & -\text{diag}(1 + \lambda_2) \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ F^H \end{bmatrix} y_0 \quad (11)$$

对式(11)求解可得 $\hat{w}^*$:

$$\hat{w}^* = \frac{\lambda_2 (\hat{x}^* \odot \hat{y}_0^*)}{\lambda_2 (\hat{x} \odot \hat{x}^*) + \lambda_1 (1 + \lambda_2)} \quad (12)$$

式(12)为权重系数离散傅里叶变换后的共轭矩阵, $\hat{x}$ 为训练样本 X 的离散傅里叶变换, $\hat{x}^*$ 为 $\hat{x}$ 的共轭矩阵, $\hat{y}_0^*$ 为目标响应离散傅里叶变换后的共轭矩阵. 对非线性分类器 $f(z)$, 将低维度输入样本 x_i 映射到高维度空间 $\varphi(x_i)$ 可以提高数据计算性能. 映射后式(12)的解可表示为输入的线性组合 $w = \sum_i \alpha_i \varphi(x_i)$, $\varphi(x_i)$ 为将训练样本 x_i 映射到高维特征空间的函数,且:

$$f(z) = w^T z = \sum_i \alpha_i \kappa(z, x_i) \quad (13)$$

其中, κ 表示核函数 $\kappa(x, x') = \varphi^T(x) \varphi(x')$. 则可得分类器系数 $\hat{\alpha}$ 为:

$$\hat{\alpha} = \frac{\lambda_2 \hat{y}_0}{\lambda_2 \hat{k}^{xx} + \lambda_1 (1 + \lambda_2)} \quad (14)$$

其中 $\hat{\alpha}$ 是 α 的离散傅里叶变换, $\hat{y}_0$ 为自适应目标响应 y_0 的离散傅里叶变换, $\hat{k}^{xx}$ 为核矩阵 K 的离散傅里叶变换.

4 抗遮挡自适应核相关滤波算法

在解决目标跟踪的遮挡问题时,一般将目标的遮挡情况分为三类^[13]: (1) 非遮挡:目标完全出现在视野内,特征信息完整;(2) 部分遮挡:目标的某一区域受到遮挡,特征信息部分缺失;(3) 完全遮挡:目标完全消失在视野内,特征信息全部缺失.

图 3 为基于实际拍摄视频采用 KCF 算法获得上述三种情况下的目标输出响应图. 其中三维图表示在检测区域中目标位置的二维高斯分布. 在 HOG 特征中, 6×6 个像素点是一个细胞, 4 个细胞是一个块, 因此响应

图中 x 轴和 y 轴表示目标大小除以每一块中所含有的细胞数得到的尺寸, z 轴表示目标输出响应值大小. 从图 3 可知在非遮挡情况下响应图中出现一个明显的尖峰, 在部分遮挡的情况下尖峰的高度有一定的减小, 而在完全遮挡时响应图趋于平滑.

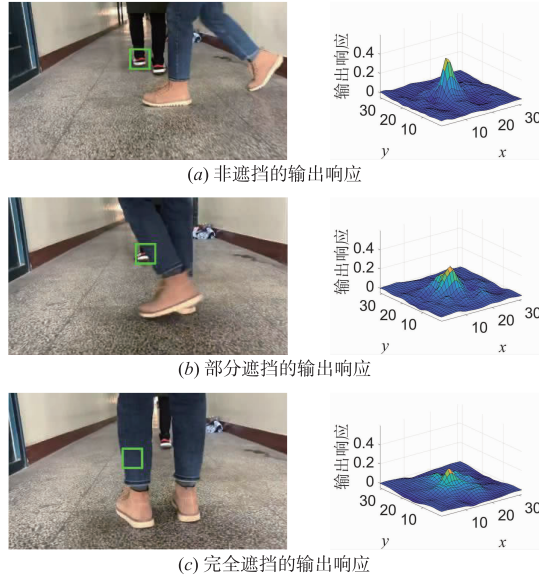


图3 不同遮挡情况下输出目标响应图

有效的遮挡检测方法对于抗遮挡跟踪的实现具有重要作用. 本文采用了一种基于响应峰值邻域相关率的遮挡检测方法. 在 KCF 算法中, 分类器检测输出目标响应值大小代表了对应样本区域与目标的相似性, 一般选择响应峰值对应的样本区域作为跟踪目标区域. 发生遮挡情况后, 足部目标受遮挡影响的区域同样会被循环移位, 输出响应峰值周围的其它响应值增大, 最终输出响应峰值变小且趋于平滑. 因此, 通过分析响应图中峰值与邻域间的数学关系可获得遮挡情况. 输出响应中峰值与邻域示意图见图 4 所示. 图 4(a) 中圆圈表示输出响应峰值, 图 4(b) 中红色区域表示峰值响应周围的其余输出响应值.

第 n 帧的相关峰值 C_p 可表示为:

$$C_p = \frac{P_n}{\frac{1}{\sigma} \sum_{t=n-\sigma+1}^n P_t} \quad (15)$$

其中, P_n 为第 n 帧的响应峰值, P_t 为第 t 帧的响应峰值, σ 为距离第 n 帧较近的帧数.

第 n 帧的相关邻域 C_N 可表示为:

$$C_N = \frac{\frac{1}{24} \left[\left(\sum_{i=x_p-5}^{x_p+5} \sum_{j=y_p-5}^{y_p+5} R_{ij} \right) - P_n \right]}{\frac{1}{\sigma} \sum_{t=n-\sigma+1}^n P_t} \quad (16)$$

其中, R_{ij} 为响应图 x - y 平面中坐标点 (i, j) 位置的响应

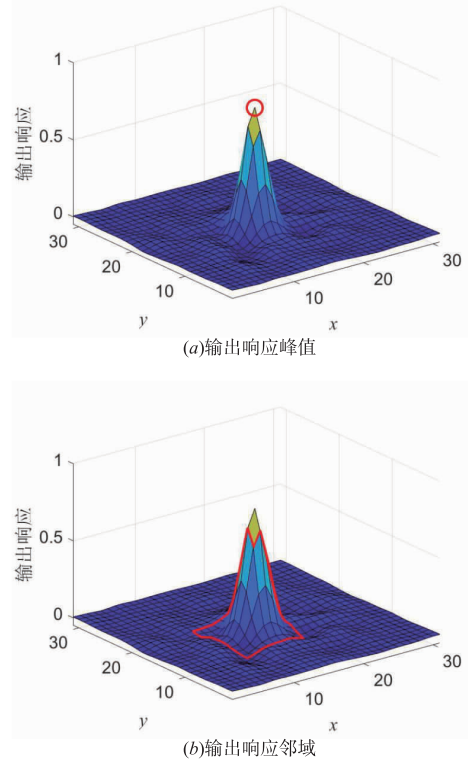


图4 输出响应峰值与邻域示意图

值, x_p 表示响应峰值处的 x 轴坐标, y_p 表示响应峰值的 y 轴坐标. 分子表示以响应峰值点坐标 (x_p, y_p) 为中心的 5×5 个单位区域中, 除去峰值点以后其余 24 个点的响应平均值.

第 n 帧输出响应的相关率 (correlation ratio) 为:

$$R_c = \frac{C_N}{C_p} \quad (17)$$

由式(17)可知, 相关率 R_c 的值越接近 1, 那么相关峰值 C_p 与相关邻域值 C_N 越相近, 当前帧的输出响应就越趋于平滑, 表明当前帧发生遮挡. 反之, R_c 的值越接近 0, 则为非遮挡状态. 因此可通过设定相关率阈值解决遮挡跟踪问题, 即当相关率大于设置阈值时, 判断当前帧为遮挡状态, 保存上一帧目标模板并停止更新模板与分类器系数.

5 运动模型引导下的自适应响应抗遮挡核相关滤波算法

运动模型引导下的自适应响应抗遮挡核相关滤波算法 (以下简称 MM-AR-AO-KCF) 算法主要在于提高跟踪准确率并解决足部跟踪中的遮挡问题. MM-AR-AO-KCF 算法步骤:

输入: 第 $t-1$ 帧图像的目标位置、分类器系数和目标观测模型.

输出: 第 t 帧图像的目标位置、分类器系数和目标

观测模型.

(1) 根据式(5)估计双脚位置,调整搜索区域,使其中心作为新的目标搜索中心;

(2) 分类器计算搜索区域内的输出响应,对响应峰值、邻域相关率进行阈值判断,若相关率大于给定阈值则为遮挡状态,停止更新分类器系数和目标观测模型,放弃对本帧处理,输入下一帧图像,并返回步骤(1),若小于给定阈值则为非遮挡状态,继续步骤(3);

(3) 找到输出响应最大位置,更新目标位置,采用上一帧的目标响应对回归标签进行修正,当回归标签与输出响应均方误差最小时,根据式(14)重新训练分类器,更新分类器系数和目标观测模型.

6 仿真实验

本文在西安理工大学校园拍摄了7组行人步行测试视频. 本文将7组视频分别命名为 Passage-way, Shade, Garden, Zoom, Turn, Occlusion2, Occlusion3, 视频情况见表1所示.

表1 实际拍摄四种行人步行视频描述

视频名称	行人行走情况和背景情况
Passage-way	室内较慢速度行走,无光照变化
Shade	室外中等速度行走,从树荫到阳光下,有明显光照变化
Garden	室外花园人行道中等步速行走,有光照变化,背景复杂
Zoom	阴天室外快速行走
Turn	室外阴天中速向左转弯
Occlusion2	室内遮挡33帧,遮挡后双脚位置较遮挡前变换较小
Occlusion3	室内遮挡63帧,遮挡后双脚位置较遮挡前变换较大

6.1 跟踪准确率

本文足部跟踪实验中算法参数设置:由文献[3]取学习率为0.02,高斯带宽为0.5. 使用自制测试集和OTB2013测试集进行多次跟踪实验得出的经验参数:正则化参数 λ_1 为 1×10^{-3} , λ_2 为 1×10^{-5} ,相关率阈值为0.75,这些参数的选取依据是获得良好的算法收敛性以及跟踪准确率.

6.1.1 运动模型引导下的自适应响应核相关滤波算法

本文针对表1中前5组非遮挡人体步行视频,采用了五种算法进行了跟踪实验,跟踪准确率对比如表2所示. 本文所提的运动模型引导的自适应输出响应(Adaptive Response, AR) KCF算法记为MM-AR-KCF,运动模型引导的KCF算法记为MM-KCF,文献[14]所提的BACF(Background Aware Correlation Filters)算法采用一个掩码矩阵增大相关滤波器的搜索样本数量,通过密

集采样的方法得到真正的负样本,该方法甚至可以在整幅图像中进行目标搜索. 文献[15]所提SAMF(Scale Adaptive kernel correlation filters with Multiple Features)算法将HOG特征与颜色特征相融合,能够有效地表征目标的颜色信息,此外针对跟踪过程中目标形变问题,采用加入尺度池的方法来提高目标跟踪的鲁棒性.

由表2可知,对5组视频跟踪的MM-KCF算法平均准确率为0.678, KCF算法的平均准确率为0.402, BACF算法的平均准确率为0.528, SAMF算法的平均准确率为0.539, 而MM-AR-KCF算法平均准确率为0.822,明显高于其它四种算法. 因此,考虑自适应输出响应对运动模型引导下的核相关滤波算法跟踪准确率有明显的提升.

表2 五种跟踪算法的跟踪准确率

视频 算法	Passage-way	Shade	Garden	Zoom	Turn
MM-AR-KCF	0.951	0.813	0.760	0.770	0.820
MM-KCF	0.882	0.765	0.720	0.710	0.312
KCF	0.862	0.392	0.140	0.320	0.298
BACF	0.921	0.716	0.430	0.290	0.282
SAMF	0.931	0.578	0.290	0.440	0.454

6.1.2 运动模型引导下的自适应响应抗遮挡核相关滤波算法

表1中有两组存在遮挡的视频,其中视频Occlusion2共100帧出现33帧遮挡, Occlusion3共120帧出现63帧遮挡. 分别采用运动模型引导下的自适应输出响应抗遮挡核相关滤波算法(MM-AR-AO-KCF)、自适应输出响应抗遮挡核相关滤波算法(AR-AO-KCF)、BACF算法以及SAMF算法对两个视频的足部跟踪准确率进行了对比实验,结果如图5所示. 准确率采用OTB评价体系的定义,即选取当前帧目标真值框中心位置和跟踪框中心位置误差为20个像素时进行计算.

由图5可知,对于视频Occlusion2, MM-AR-AO-KCF算法的准确率为0.860, AR-AO-KCF算法的准确率为0.83, BACF算法的准确率为0.400, SAMF算法的准确率为0.390. 对视频Occlusion3, MM-AR-AO-KCF算法的准确率为0.775, AR-AO-KCF算法的准确率为0.167, BACF算法的准确率为0.141, SAMF算法的准确率为0.158. 可以看出,在遮挡帧数较多时,运动模型引导下的自适应输出响应抗遮挡核相关滤波算法可以有效解决目标遮挡问题.

6.2 跟踪误差

跟踪误差定义为每帧的跟踪框的中心与真值框的

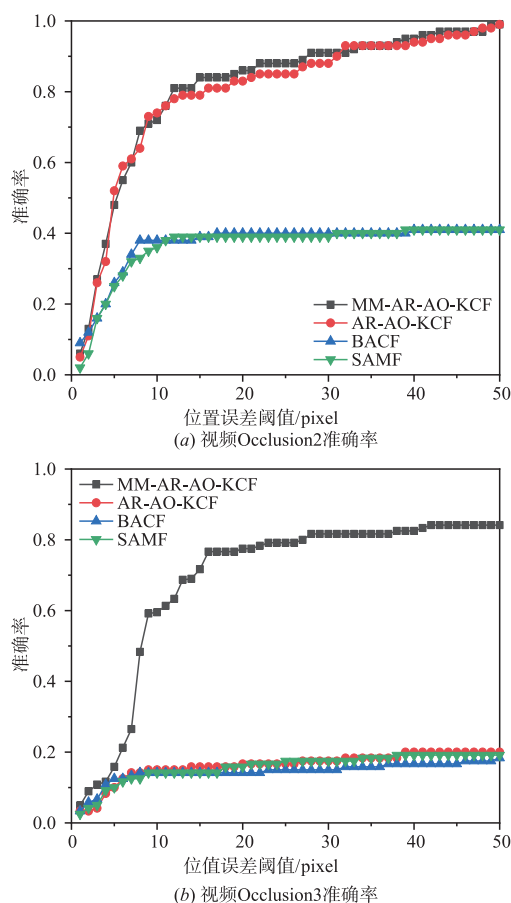


图5 两个有遮挡视频的跟踪准确率

中心坐标之间的欧式距离,单位为像素.图6为采用4种算法处理的两组有遮挡视频的跟踪误差曲线图.

由图6(a)可知,在第40帧遮挡开始后,SAMF算法和BACF算法跟踪框漂移,跟踪误差迅速增大,而MM-AR-AO-KCF算法和AR-AO-KCF算法采用了自适应输出响应和抗遮挡方法,在遮挡结束后这两种算法仍继续成功跟踪目标.由图6(b)可知,该视频受遮挡的帧数多,在20帧遮挡开始后,BACF算法和SAMF算法的跟踪误差增大,AR-AO-KCF算法虽然停止了滤波器系数更新保留原有位置,但是由于遮挡结束后双脚位置超出了检测区域,所以在遮挡结束后跟踪误差依然较大.而MM-AR-AO-KCF算法因为运动模型的位置预测使目标始终包含在内检测区域,跟踪误差较小,成功跟踪目标.

6.3 跟踪结果

对于拐弯视频 Turn 选取表2中MM-AR-KCF算法和SAMF算法的跟踪实验结果进行对比,以第10、60、110、170帧为例.而对于遮挡视频 Occlusion2 和视频 Occlusion3,因为SAMF算法在遮挡情况下的跟踪性能不佳,因此本节对视频 Occlusion2 和视频 Occlusion3 下采用MM-AR-AO-KCF算法和AR-AO-KCF算法的实验

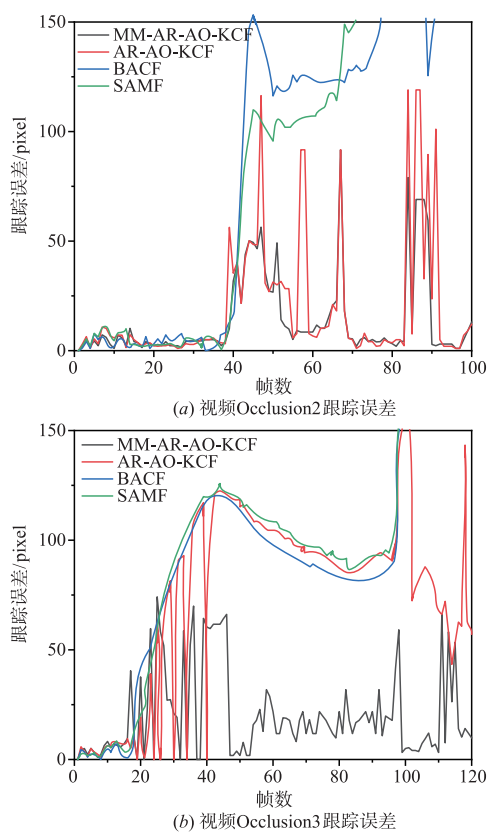


图6 两个有遮挡视频的跟踪误差

结果进行对比.两组视频的时长不同,视频 Occlusion2 的跟踪结果选取以第1、33、66、99帧为例,视频 Occlusion3 的跟踪结果选取以第1、40、80、120帧为例,跟踪结果见图7所示.

对比图7(a)和图7(b)可以看出在跟踪开始时直线行走两种算法的跟踪效果相差不大,但是在转弯的过程中,由于双脚特征的相似性,SAMF算法的左脚跟跟踪目标切换到了右脚上,导致了跟踪的失败,而MM-AR-KCF算法始终成功跟踪双脚目标.对比图7(c)和图7(d)可以看出,两种算法对视频 Occlusion2 的跟踪效果相近,两种算法在38帧遮挡开始后,均停止更新目标外观模型和滤波器系数,防止遮挡带来的影响,在本视频遮挡前后双脚在图像中的位置变化较小,在遮挡结束后,双脚依旧处于检测区域内,所以在跟踪效果上MM-AR-AO-KCF算法与AR-AO-KCF算法均成功跟踪.视频 Occlusion3 相对于视频 Occlusion2 遮挡时间更长,且遮挡结束后双脚较遮挡前的位置变化较大.对比图7(e)和图7(f)可以看出,AR-AO-KCF算法在遮挡结束后丢失了目标,这是因为足部已经超出了跟踪算法检测区域的范围,导致跟踪目标必然丢失.而MM-AR-AO-KCF算法采用运动模型首先预估出下一帧的检测区域位置,再完成模板匹配,因此在遮挡结束后,足部目标依然在检测区域内,最后成功跟随目标.



图7 有遮挡情况的视频跟踪结果图

7 运动模型引导下的自适应抗遮挡核相关滤波算法机器人试验

本文采用 Turtlebot 机器人和 ROS (Robot Operating System)^[16] 完成试验设计. 机器人上搭载了微软 Kinect V1 视觉传感器和酷睿 i5 处理器、频率 2.4GHz 笔记本. 安装 Ubuntu16.04 版本的操作系统和 ROS Kinetic 版本. 机器人的参数设置如下: 最大跟随距离 5m, 最小跟随距离 1.5m, 最大线速度 0.6m/s, 最小线速度 0.3m/s, 最大角速度 0.7rad/s. 通过 Kinect 判断双脚与机器人之间的距离以及跟踪算法所预估双脚的位置, 得到机器人移动所需的线速度与角速度. 在室内走廊进行了真实场景下人体目标跟踪实验. 实验场所为西安理工大学教 5 楼走廊, 光照明亮、稳定, 目标人以较慢步速行走, 行走过程中从左侧有人穿过, 对目标人的双脚进行遮挡.

7.1 较短时间遮挡

机器人跟踪试验中行人从左侧穿过未做停留, 对



图8 短暂遮挡情况跟踪算法结果图

目标人双脚遮挡 27 帧, 彩色摄像头帧率 30FPS, 遮挡时间为 0.90s. 在目标人的双脚被遮挡过程中, 跟踪框有一定的漂移, 如图 8 所示, 但在遮挡结束后能够继续跟踪双脚, 完成机器人的跟随任务, 如图 9 所示. 本试验验证了 MM-AR-AO-KCF 算法能克服短暂遮挡对目标跟踪带来的影响.

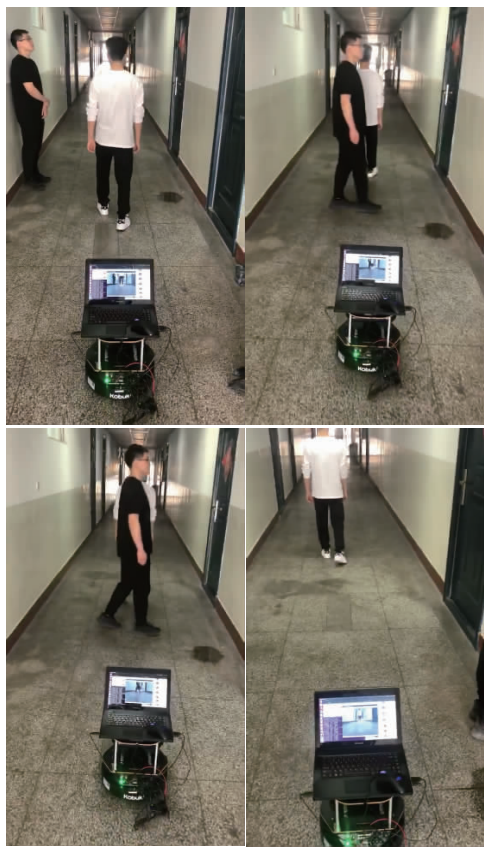


图9 短暂遮挡情况机器人跟随结果图

7.2 较长时间遮挡

试验中行人从左侧走入视野中,做了一定时间的驻足后迈步离开,对目标人的双脚遮挡了74帧,遮挡时间为2.47s.目标人的双脚被完全遮挡后,跟踪框发生了漂移,如图10所示.根据MM-AR-AO-KCF算法原理,遮挡时目标模板和分类器系数停止更新,保留无遮挡时的目标模板和分类器系数,在遮挡结束后成功检测到双脚,机器人成功跟随目标人,如图11所示.本试验验证了即使在较长时间遮挡的情况下,MM-AR-AO-KCF算法依然能够成功跟踪目标,使机器人能够稳定的跟随目标人.

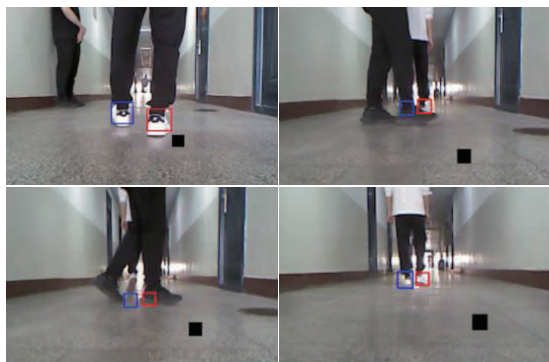


图10 较长时间遮挡情况机器人跟随结果图



图11 较长遮挡情况机器人跟随结果图

8 结论

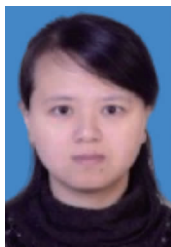
本文针对小型移动机器人高度低的特点,将人体双脚作为跟踪目标.通过建立人体足部运动模型,对下一帧图像双脚位置进行预测,再由预测位置信息获得核相关滤波算法的目标搜索区域,提高了跟踪准确性与鲁棒性,能够解决在跟踪过程中遇到的光照变化、快速步行等较复杂问题.同时,针对核相关滤波算法存在的问题,提出了一种自适应输出响应抗遮挡跟踪算法,进一步提高了跟踪准确率,并成功解决了跟踪中的遮挡问题.此外,本文基于ROS系统的移动机器人,在现实场景进行了遮挡场景下的跟踪试验,验证了所提算法应用于人体足部跟踪具有良好的鲁棒性和实时性.

参考文献

- [1] Bolme D S, Beveridge J R, Draper A B. Visual object tracking using adaptive correlation filters [A]. Twenty-third IEEE Conference on Computer Vision & Pattern Recognition [C]. Los Alamitos: IEEE Computer Society Press, 2010. 2554 – 2560.
- [2] Henriques J F, Caseiro R, Martins P. High-speed tracking with kernelized correlation filters [J]. IEEE Transactions

- on Pattern Analysis and Machine Intelligence, 2014, 37(3):583–596.
- [3] Wu Y, Lim J, Yang M H. Object tracking benchmark[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2015, 37(9):1834–1848.
- [4] Kumara R, Maria A. Object tracking with adaptive motion modeling of particle filter and support vector machines[A]. 2015 IEEE International Conference on Image Processing[C]. Los Alamitos: IEEE Computer Society Press, 2015. 353–360.
- [5] Wang X P, Zhao Q J, Tan H K. Self-tuning visual tracking based on multiple motion models[A]. 2016 35th Chinese Control Conference[C]. Los Alamitos: IEEE Computer Society Press, 2016. 3990–3994.
- [6] Wang Y, Wang H J, Li C. Methods of target motion estimation for AUV target tracking[A]. 2016 IEEE International Conference on Mechatronics and Automation[C]. Los Alamitos: IEEE Computer Society Press, 2016. 2139–2144.
- [7] 陈婧, 孙玉娟, 周万军. 融合运动模型与联合置信度量的改进核相关跟踪算法[J]. 红外技术, 2018, 40(11):1106–1111.
Chen J, Sun Y J, Zhou W J. An improved kernelized correlation tracking algorithm based on a joint confidence measurement and motion model[J]. Infrared Technology, 2018, 40(11):1106–1111. (in Chinese)
- [8] 廖家才, 曹立波, 夏家豪. 基于核相关滤波和运动模型的多目标轨迹跟踪[J]. 汽车工程, 2019, 41(10):1179–1188.
Liao J C, Cao L B, Xia J H. Multi-target trajectory tracking based on kernelized correlation filtering and motion model[J]. Automotive Engineering, 2019, 41(10):1179–1188. (in Chinese)
- [9] 程熙. 基于深度图像的室内移动机器人行人跟随[D]. 长沙: 湖南大学, 2018.
Chen X. Pedestrian Following Method for Indoor Mobile Robot Based on Depth Image[D]. Changsha: Hunan University, 2018. (in Chinese)
- [10] Li Y, Ding S, Zhai Q. Human feet tracking guided by locomotion model[A]. 2015 International Conference on Robotics & Automation[C]. Los Alamitos: IEEE Computer Society Press, 2015. 2424–2429.
- [11] 闫培亮. 基于相关滤波的目标跟踪算法研究[D]. 西安: 西安电子科技大学, 2019.
Yan P L. Research on Object Tracking Algorithm Based on Correlation Filtering[D]. Xi'an: Xidian University, 2019. (in Chinese)
- [12] Kiani G H, Sim T, Lucey S. Correlation filters with limited boundaries[A]. IEEE Conference on Computer Vision and Pattern Recognition[C]. Los Alamitos: IEEE Computer Society Press, 2015. 4630–4638.
- [13] 刘海峰, 孙成, 梁星亮. 自适应特征融合与抗遮挡的相关滤波跟踪算法[J/OL]. 激光与光电子学进展, <http://kns.cnki.net/kcms/detail/31.1690.TN.20191217.1448.026.html>, 2019-12-11.
Liu H F, Sun C, Liang X L. Correlation filter tracking algorithm with adaptive feature fusion and anti-occlusion[J/OL]. Laser & Optoelectronics Progress, <http://kns.cnki.net/kcms/detail/31.1690.TN.20191217.1448.026.html>, 2019-12-11. (in Chinese)
- [14] Galoogahi H K, Fagg A, Lucey S. Learning background-aware correlation filter for visual tracking[A]. 2017 International Conference on Computer Vision[C]. Los Alamitos: IEEE Computer Society Press, 2017. 1144–1152.
- [15] Li Y, Zhu J. A scale adaptive kernel correlation filter tracker with feature integration[A]. European Conference on Computer Vision[C]. Heidelberg: Springer Press, 2015. 254–265.
- [16] Lentin Joseph. Mastering ROS for Robotics Programming[M]. Birmingham: Packt Publishing, 2015.

作者简介



陈丹 女, 1975 年生于陕西汉中. 博士, 硕导. 现为西安理工大学自动化与信息工程学院副教授. 主要研究方向为现代信号处理, 机器人目标识别与跟踪.
E-mail: chdh@xaut.edu.cn



姚伯羽 男, 1994 年生于陕西西安. 现为西安理工大学自动化与信息工程学院硕士研究生. 主要研究方向为计算机视觉.
E-mail: 423058934@qq.com