

基于卡尔曼滤波的多目标跟踪方法

赵广辉 卓 松 徐晓龙

(武汉理工大学计算机科学与技术学院 武汉 430070)

摘 要 针对视频多目标跟踪中由于目标间的遮挡、交错或目标漂移而导致跟踪失败的情况,提出一种基于卡尔曼滤波以及空间颜色直方图的遮挡预测跟踪算法。利用空间颜色直方图对目标进行建模,可以对不同目标进行区分进而在目标之间出现交错或目标漂移时仍能跟踪到目标。通过卡尔曼滤波算法可以预测目标的状态,对预测位置之间存在交错的目标进行遮挡标记,以便在下一帧中仍然可以跟踪到被遮挡的目标。采用 2D MOT 2015 数据集进行实验,跟踪的平均精度达到了 34.1%。实验结果表明,所提方法对多目标跟踪的效果有所提高。

关键词 多目标跟踪,卡尔曼滤波,空间颜色直方图,遮挡预测

中图分类号 TP391.41 文献标识码 A DOI 10.11896/j.issn.1002-137X.2018.08.045

Multi-object Tracking Algorithm Based on Kalman Filter

ZHAO Guang-hui ZHUO Song XU Xiao-long

(School of Computer Science and Technology, Wuhan University of Technology, Wuhan 430070, China)

Abstract Aiming at the tracking failure caused by occlusion between objects, interleaving or target drift in multi-object tracking, this paper proposed an occlusion prediction tracking algorithm based on Kalman filter and spatiograms. By combining the color histogram and the distribution of color in space, spatiograms can be used to distinguish different objects, so that the object can still be tracked when interleaving or occlusion between objects occurs. The state of the object can be predicted by the Kalman filtering algorithm. The occlusion mark is used for the object which overlaps with other objects, so that the occluded object which is undetected can be tracked in the next frame. The 2D MOT 2015 data set was used for experiment. The average accuracy of tracking achieves 34.1%. Experimental results show that the algorithm can improve the performance of multi-object tracking.

Keywords Multi-object tracking, Kalman filter, Spatial color histogram, Occlusion prediction

1 引言

视频中的多目标跟踪是计算机视觉领域的一个重要研究方向。近年来,得益于目标检测技术的发展^[1-2],基于目标检测的多目标跟踪方法成为当前多目标跟踪的主要方法。基于检测的多目标跟踪方法的跟踪过程能够被转化为数据关联问题,即将已经跟踪到的多个目标和检测到的多个目标关联起来。显然,基于目标检测的方法对检测结果的依赖程度较高,复杂背景和目标间的遮挡会对目标检测造成较大影响,进而给数据关联造成影响;同时在多个检测值和多个跟踪值之间建立准确的对应关系所依赖的目标模型也会较大程度地影响多目标跟踪的效果。

根据在跟踪目标和检测目标间建立的关联所依据的检测集的不同,可以将多目标跟踪分为离线跟踪和在线跟踪两类。对于离线的多目标跟踪,可以同时使用当前的检测信息和后面的检测信息进行运动轨迹优化来实现跟踪^[3-6],这类多目标跟踪方法利用后面的检测信息解决目标漏检、目标遮挡以及

目标漂移的问题,但是这类方法因为考虑了大量检测信息来进行轨迹优化而增加了计算量,在一些实时性要求较高的场景中不具有应用价值,因此需要使用具有更高效率的在线跟踪方法。在线跟踪方法只根据被跟踪目标信息和当前帧检测的信息来实现数据关联,在线跟踪通常使用基于概率的算法来解决因为目标漏检和目标漂移等问题造成的不确定性^[7-10]。为了保证匹配过程的准确性,一些方法对目标的运动过程或者目标的外观进行建模^[11-12]。

解决多目标跟踪中的数据关联问题的方法有多假设跟踪方法^[5]、关联滤波器方法^[6]、马尔可夫链蒙特卡罗方法^[13],这些方法的特点是为了应对跟踪中的不确定性而维持多个假设,但是这样做会导致计算量增大,且随着跟踪目标个数的增多,效率会显著下降。匈牙利算法^[14]以及贪婪关联方法^[15]是解决多目标跟踪中数据关联问题更为高效的方法。Bewley等^[9]用匈牙利算法来关联被跟踪目标和检测目标,在保证效率的同时取得了较好的跟踪结果。Breitenstein等^[15]采用粒子滤波作为预测和更新目标状态的方法,并使用贪婪关联方

到稿日期:2017-12-30 返修日期:2018-03-08 本文受中央高校基本科研业务费项目(2017-zy-084)资助。

赵广辉(1973—),男,博士,副教授,主要研究方向为智能计算、机器学习,E-mail:zhao@whut.edu.cn(通信作者);卓松(1991—),男,硕士生,主要研究方向为计算机视觉、视频目标跟踪;徐晓龙(1995—),男,硕士生,主要研究方向为图像处理、计算机视觉。

法进行数据关联,利用目标检测过程中的置信度作为判断遮挡的依据。袁大龙等^[16]使用匈牙利算法来解决数据关联问题,并对未遮挡目标和被遮挡目标分别采用不同的预测和更新方法,提高了目标跟踪的适应性。

目标的模型作为数据关联的依据也会影响跟踪效果,Birchfield等^[17]在传统的颜色直方图的基础上加入了像素位置信息,从而形成了空间颜色直方图,丰富了目标外观模型的信息,使得跟踪更加鲁棒。Conaire等^[18]在空间颜色直方图的基础上提出了新的计算目标相似度的方法。Yang等^[19]提出了基于多部分的人体模型在线学习方法来对目标进行建模,该模型可以较好地解决目标被部分遮挡时的跟踪问题,但是该模型的复杂性降低了跟踪效率。

对于基于目标检测的在线多目标跟踪框架,Bewley等^[9]提出了一种基于卡尔曼滤波^[20]的跟踪方法 SORT,该方法使用观测目标与预测目标间的交叠率来构建关联矩阵,并使用匈牙利算法进行数据关联。这种方法的效率较高且能够得到较好的跟踪效果,但是由于仅仅使用了预测目标和观测目标之间的交叠率作为数据关联的唯一线索,易出现目标交错、目标漂移导致的跟踪失败问题,且该方法对目标检测的效果较为敏感。与之相对的是本文使用自适应尺度空间颜色直方图对目标进行建模,有效地减少了目标交换和目标漂移造成的跟踪失败的情况,并且利用跟踪过程中目标之间的位置关系对目标间的遮挡进行预测,从而在漏检时结合数据关联状态对漏检的目标实现跟踪。

2 多目标跟踪框架

2.1 卡尔曼滤波

卡尔曼滤波是一种最小化均方误差的线性滤波方法,它在观测方程和目标状态转移方程的基础上,使用迭代的方式对目标的状态进行预测和更新。粒子滤波也是目标跟踪问题中较为常用的预测和更新方法,相对于卡尔曼滤波方法,粒子滤波方法需要为每个目标维护多个状态并且预测和更新过程较为耗时,因此在实时性要求较高的场景中使用卡尔曼滤波方法更为合适。卡尔曼滤波的不足之处在于对目标运动速度的改变不具备较好的适应性,对于此问题,本文使用较为准确的目标检测算法以及改进的空间颜色直方图来弥补目标位置预测的不准确性。状态转移方程和观测方程是卡尔曼滤波的基础,其公式如下:

$$x_t = A_{t|t-1} x_{t-1} + w_{t-1} \quad (1)$$

$$z_t = H x_t + v_t \quad (2)$$

其中, z_t 表示 t 时刻系统的观测状态, x_t 和 x_{t-1} 对应 t 时刻和 $t-1$ 时刻系统的状态。 $A_{t|t-1}$ 表示状态转移矩阵,对应目标的运动方式; H 是观测矩阵,对应内部状态和观测值之间的关系。 v_t 对应测量噪声,服从正态分布,其协方差为 R ; w_t 对应系统噪声,同样满足正态分布,其协方差为 Q 。预测和更新方程如下:

$$\bar{x}_{t|t-1} = A_{t|t-1} x_{t-1}^u \quad (3)$$

$$P_{t|t-1} = A_{t|t-1} P_{t-1} A_{t|t-1}^T + Q_{t-1} \quad (4)$$

$$K_t = P_{t|t-1} H^T (H P_{t|t-1} H^T + R_t)^{-1} \quad (5)$$

$$x_t^u = \bar{x}_{t|t-1} + K_t (z_t - H \bar{x}_{t|t-1}) \quad (6)$$

$$P_t = P_{t|t-1} - K_t H P_{t|t-1} \quad (7)$$

其中,式(3)表示状态预测方式, x_{t-1}^u 对应 $t-1$ 时刻的最优估计状态, $\bar{x}_{t|t-1}$ 是在最优估计基础上的预测状态;式(4)表示误差协方差的预测方式, P_{t-1} 对应 $t-1$ 时刻更新后的误差协方差矩阵, $P_{t|t-1}$ 表示误差协方差的预测值;式(5)表示卡尔曼增益计算方式, K_t 对应 t 时刻的卡尔曼增益值,用于后续对目标状态预测值和误差协方差预测值进行优化;式(6)表示目标状态优化方式, x_t^u 对应 t 时刻的最优预测值;式(7)表示误差协方差的优化方式, P_t 对应 t 时刻的最优估计值。

一般情况下,同一个目标在相邻帧之间的运动方式可以近似为匀速直线运动,那么目标的位置 L 、速度 V 、目标矩形框面积 S 及面积变化率 V_s 和间隔时间 Δt 之间有如下关系:

$$L(t) = L(t-1) + \Delta t V(t-1) \quad (8)$$

$$S(t) = S(t-1) + \Delta t V_s(t-1) \quad (9)$$

$$V(t) = V(t-1) \quad (10)$$

$$V_s(t) = V_s(t-1) \quad (11)$$

状态向量 x 需要记录位置、速度以及尺寸信息,可以表示为:

$$x = [l_x \ l_y \ s \ r \ v_x \ v_y \ v_s]^T \quad (12)$$

其中, l_x 和 l_y 对应目标中心在两个坐标轴上的投影分量, v_x 和 v_y 对应目标运动速度在两个坐标轴方向上的分量, r 对应宽度和高度之比, s 表示面积, v_s 表示面积变化率。结合上面的关系方程,规定 Δt 为单位时间,状态转移矩阵可以表示如下:

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad (13)$$

由于状态向量中的位置和尺寸可以直接作为观测向量中的值,则观测矩阵表示如下:

$$H = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix} \quad (14)$$

系统噪声协方差矩阵 Q 和观测噪声协方差矩阵 R 对应的值如下所示:

$$Q = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad (15)$$

$$R = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (16)$$

在卡尔曼滤波器的初始化阶段,目标的位置和面积可以在目标检测过程中得到,将目标的速度和面积变化率初始化为 0,考虑到初始化造成的随机性较大,对初始的协方差矩阵赋值如下:

$$P_0 = \begin{pmatrix} 100 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 100 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 100 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 100 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 100 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 100 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 100 \end{pmatrix} \quad (17)$$

卡尔曼滤波器初始化之后就可以根据后续的目标检测信息在一个迭代过程中对目标的状态进行预测和最优化估计,该迭代过程即对应目标跟踪过程。

2.2 空间颜色直方图

2.2.1 尺度不变的空间颜色直方图

空间颜色直方图^[17]在传统颜色分布直方图的基础上加入了颜色分布信息。传统的颜色分布直方图对灰度区间的数量分布进行统计,空间颜色直方图不仅涵盖了数量统计信息,还统计了每个灰度区间的空间分布(包括空间均值和协方差矩阵)。相比于传统颜色直方图,空间颜色直方图丰富了目标的外观描述,对目标有更强的辨识能力,进而增加了数据关联的准确性。

空间颜色直方图包含灰度区间的数量统计信息、空间均值和协方差矩阵,如式(18)所示:

$$h_i^{(2)}(b) = \langle n_b, \mu_b, \Sigma_b \rangle, b=1, \dots, B \quad (18)$$

其中, n_b 对应灰度区间 b 中的像素点占有所有像素点的比例, μ_b 对应区间 b 中所有像素点坐标的均值, Σ_b 对应区间 b 中所有像素点坐标的协方差矩阵, B 对应灰度区间的数量。可以发现,如果忽略均值 μ_b 和协方差 Σ_b ,空间颜色直方图就退化成传统的颜色直方图。

为了利用空间颜色直方图进行数据关联,需要计算两个目标对应的空间直方图的相似度,计算公式如下:

$$\rho(h, h') = \sum_{b=1}^B w_b \rho_b(n_b, n_b') \quad (19)$$

w_b 对应不同灰度区间的加权系数,体现了对应灰度区间对整体的影响大小,该系数用高斯函数来描述,如下所示:

$$w_b = \eta \exp\left[-\frac{1}{2}(\mu_b - \mu_b')^T \hat{\Sigma}_b^{-1}(\mu_b - \mu_b')\right] \quad (20)$$

$$\hat{\Sigma}_b^{-1}(y) = \Sigma_b^{-1}(y) + (\Sigma_b')^{-1} \quad (21)$$

其中, η 表示高斯归一化系数。 $\rho_b(n_b, n_b')$ 使用 Bhattacharyya 相似度来描述两个颜色直方图的相似度, $\rho_b(n_b, n_b') = \sqrt{n_b n_b'}$ 。

空间颜色直方图中灰度区间 b 的坐标均值向量 μ_b 和协方差 Σ_b 表示如下:

$$\mu_b = \frac{1}{\sum_{j=1}^{N_b} \delta_{jb}} \sum_{i=1}^{N_b} (x_i - y) \delta_{jb} \quad (22)$$

$$\Sigma_b = \frac{1}{\sum_{j=1}^{N_b} \delta_{jb}} \sum_{i=1}^{N_b} (x_i - \mu_b(y))(x_i - \mu_b(y))^T \delta_{jb} \quad (23)$$

其中, y 对应目标中心所在位置的坐标向量, x_i 对应目标矩形

区域中各像素点的坐标向量。 δ_{jb} 根据像素点 x_i 的灰度值是否等于 b 区间代表的灰度值来取值。如果相等,则 δ_{jb} 的值为 1;如果不相等,则 δ_{jb} 的值为 0。通过计算两个目标对应的空间颜色直方图的相似度,可以度量两个目标的关联程度,为目标关联提供依据。

2.2.2 适应尺度的空间颜色直方图

空间颜色直方图被提出时针对的是尺度不变的目标,没有考虑目标在运动过程中尺度的改变,由于目标尺度在跟踪过程中变化的可能性较大,如果直接应用式(22)和式(23)来计算灰度区间的均值和协方差会存在较大误差,因此,本文加入尺度作为修正变量,以修正灰度区间的坐标均值和协方差矩阵,使二者在目标尺度改变的情况下能够反映颜色空间分布的不变性。尺度信息在目标检测时获得,并在跟踪过程中被保存、更新。修正后的坐标均值向量和协方差矩阵表示如下:

$$\mu_b = \frac{1}{\sum_{j=1}^{N_b} \delta_{jb}} \sum_{i=1}^{N_b} \delta_{jb} s(x_i - y) \quad (24)$$

$$\Sigma_b = \frac{1}{\sum_{j=1}^{N_b} \delta_{jb}} \sum_{i=1}^{N_b} \delta_{jb} s(x_i - \mu_b(y))(x_i - \mu_b(y))^T s^T \quad (25)$$

$$s = \begin{pmatrix} \frac{1}{c} & 0 \\ 0 & \frac{1}{d} \end{pmatrix} \quad (26)$$

其中, c 对应目标矩形框的宽度, d 对应其高度。将目标尺度作为参考因素能够反映出颜色在空间分布上的不变性,可以提高目标间的关联的准确性。利用原始的空间颜色直方图模型和自适应尺度的空间颜色直方图模型计算图 1(a)和图 1(b)中同一目标的相似度,使用原始的空间颜色直方图计算所得的相似度为 0.62,使用改进的方法计算所得的相似度为 0.83。如果在跟踪过程中不考虑目标尺度的改变而直接应用原始的空间颜色直方图对目标进行建模,将会导致相似性计算值存在误差。通过将目标尺度当作修正变量,自适应尺度的空间颜色直方图能够反映颜色在空间分布上的不变性,在跟踪过程中能够更加准确地判断目标间的关联程度,为目标关联提供依据。



(a) (b)
注:目标用方框标记

图 1 不同尺寸大小的同一目标

Fig. 1 Same object with different sizes

2.2.3 基于融合特征的关联矩阵

对于基于检测的多目标跟踪,跟踪的关键过程是在跟踪目标和检测目标之间建立一对一关系。本文结合跟踪目标和检测目标间的自适应空间颜色直方图相似度和交叠率来构造关联矩阵,然后使用匈牙利算法计算跟踪目标和检测目标间的对应关系。假设当前被跟踪目标的数量为 n ,检测目标数量为 m ,根据空间颜色直方图相似度的计算方法可以计算跟

踪目标和预测目标之间的自适应空间颜色直方图相似度矩阵 H :

$$H = \begin{pmatrix} \rho_{11} & \cdots & \rho_{1n} \\ \vdots & \ddots & \vdots \\ \rho_{m1} & \cdots & \rho_{mn} \end{pmatrix} \quad (27)$$

其中, ρ_{ij} 表示第 i ($i=1, \dots, m$) 个检测目标和第 j ($j=1, \dots, n$) 个被跟踪目标间的自适应空间颜色直方图相似度。自适应尺度的空间颜色直方图相似度体现的是跟踪目标和检测目标在外观上的关联程度, 交叠率则体现了跟踪目标和检测目标在空间位置上的关联程度, 通过交叠率可以构建基于位置的相似度矩阵 P 。交叠率 iou 以及矩阵 P 的计算公式如下:

$$iou_{ij} = \frac{area_m}{area_i + area_j - area_m} \quad (28)$$

$$P = \begin{pmatrix} iou_{11} & \cdots & iou_{1n} \\ \vdots & iou_{ij} & \vdots \\ iou_{m1} & \cdots & iou_{mn} \end{pmatrix} \quad (29)$$

其中, $area_i$ 表示第 i 个检测目标矩形框 R_i 的面积; $area_j$ 表示第 j 个被跟踪目标矩形框 R_j 的面积; $area_m$ 表示 R_i 和 R_j 这两个矩形重叠区域的面积; iou_{ij} 表示第 i 个检测目标与第 j 个被跟踪目标间的交叠率。

通过结合基于目标外观模型关联程度的关联矩阵 H 以及基于位置关联程度的关联矩阵 P , 可以得到鲁棒性更好的融合关联矩阵 M , M 的计算方法如下:

$$M = \alpha H + (1 - \alpha) P \quad (30)$$

实数 α 表示融合系数, 取值范围在 0 到 1 之间。相比 SORT 方法仅采用目标间的交叠率构建关联矩阵, 本文使用融合特征建立关联矩阵, 对不同场景有更好的适应性。如图 2 和图 3 所示, 当目标交错时, SORT 跟踪结果出现了目标交换错误, 本文的方法能够准确区分两个交错的目标(融合系数 α 取值为 0.4)。

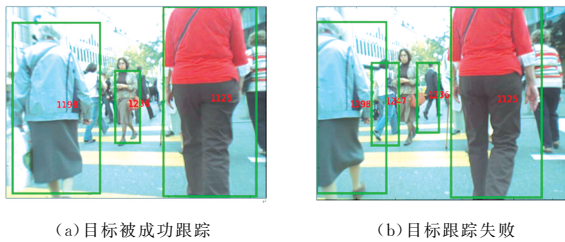


图 2 SORT 方法出现的跟踪错误
Fig. 2 Tracking error of SORT method



图 3 本文方法的跟踪效果
Fig. 3 Tracking results of proposed method

2.3 遮挡预测

对于基于检测的多目标跟踪方法, 跟踪过程对检测结果的依赖程度较大, 检测结果的好坏直接影响后续的跟踪效果。

在目标检测中经常出现由于目标间的遮挡导致目标漏检进而导致跟踪失败的情况, 因此跟踪过程中必须将目标间的遮挡作为考虑因素。为了降低由于目标间的遮挡造成的目标漏检对后续跟踪过程的影响, 本文提出了遮挡预测方法。遮挡预测方法利用前面跟踪过程中的目标空间分布信息来对目标间可能会出现遮挡进行预测, 如果目标预测位置和其他目标预测位置有交叠区域, 则对目标进行标记, 以便在后续的处理过程中迅速查找目标是否被遮挡。在对后续的跟踪目标和检测目标建立关联的过程中, 如果某个已经被标记过的目标没有匹配目标, 且该目标位置不处于整个画面的边缘处, 则将目标位置预测值作为目标位置。为了减少由于目标在较长时间内被遮挡而逐渐累积的预测误差, 本文规定在无目标匹配情况下被标记目标位置的最大预测次数为 2。

本文提出的遮挡预测方法对于由目标间的遮挡而造成目标漏检情况的敏感性较低, 在目标间出现遮挡的情况下仍然能够完成跟踪。如图 4 所示, 在不应用遮挡预测方法时, 2 号目标由于被其他目标遮挡而漏检, 从而造成跟踪失败。图 5 表明, 遮挡预测方法在同样的情况下完成了目标跟踪。



(a) 2 号目标被成功跟踪 (b) 2 号目标跟踪失败

图 4 由于目标被遮挡而导致的跟踪失败

Fig. 4 Tracking failure because of object occlusion



(a) (b)

图 5 遮挡预测方法的跟踪效果

Fig. 5 Tracking results of occlusion prediction method

2.4 算法流程

基本跟踪框架主要包括目标检测、卡尔曼滤波预测、检测值和预测值间的匹配以及遮挡预测, 检测算法是 FrRCNN。与 SORT 方法所采用的仅仅使用检测目标和预测目标间的交叠率作为检测值和预测值匹配的唯一依据不同, 本文提出的方法不仅使用交叠率作为匹配依据, 同时加入了预测目标和检测目标间的空间颜色直方图的相似度作为匹配依据, 通过二者的结合来提高关联准确性, 从而有效减少如目标交换和目标漂移丢失目标的情况。通过遮挡预测可以在一定程度上减小漏检对跟踪造成的影响。本文提出的多目标跟踪方法的流程如下。

- 1) 读取图像序列中的第一帧图像, 并使用 FrRCNN 方法进行目标检测;
- 2) 对于图像中的每个检测目标, 初始化卡尔曼滤波器, 并使用自适应尺度空间颜色直方图对目标进行建模;

3)应用卡尔曼滤波方法预测目标的状态,并进行遮挡预测;

4)读入下一帧图像,进行目标检测,计算跟踪目标和检测目标间的自适应尺度颜色直方图相似度关联矩阵 H ,计算跟踪目标和检测目标间的交叠率关联矩阵 P ,融合关联矩阵 H 和 P 形成融合关联矩阵 M ,并使用匈牙利算法进行目标关联;

5)根据遮挡预测的结果,对未匹配的被标记的目标设定其假想的检测位置;

6)结合目标预测状态和检测信息,利用卡尔曼滤波方法对目标的状态进行更新,然后转到步骤 3)进行后续跟踪过程,直到所有图像已被处理完。

3 实验结果与分析

实验所采用的数据集为 2D MOT 2015,该数据集包含 11 组不同场景的图像集,包含固定摄像头和移动摄像头两种拍摄方式。实验平台为:处理器 Intel i5-4210U,内存 4GB,软件工具 Python2.7。

表 1 是本文方法和对比方法 SORT 在图像集 PETS09-S2L1 上的实验结果,该集合为固定摄像头场景,图像数量为 795,目标数量为 4476,路径数为 19。表 2 是在 KITTI-13 图像集上的实验结果,该集合为移动摄像头场景,图像数量为 340,目标数量为 762,路径数量为 42。

表 1 PETS09-S2L1 图像集上的对比结果

Table 1 Comparison results on PETS09-S2L1

Methods	MOTA ↑	MT ↑	FP ↓	FN ↓	ID sw ↓
SORT ^[9]	61.9	8	485	1119	102
Proposed	63.7	12	503	823	93

表 2 KITTI-13 图像集上的对比结果

Table 2 Comparison results on KITTI-13

Methods	MOTA ↑	MT ↑	FP ↓	FN ↓	ID sw ↓
SORT ^[9]	19.3	0	131	468	16
Proposed	21.7	4	115	437	11

结合表 1、表 2 的实验结果可以发现,本文方法改善了多目标跟踪的效果。对于 PETS09-S2L1,其目标数量较大,目标间遮挡、交错情况较多,SORT 方法仅采用交叠率作为关联依据,当目标间距离较近时易出现关联错误;而本文采用相似度融合的方式有效地减少了关联错误,遮挡预测方法减少了由于目标间遮挡导致的跟踪失败的情况。对于 KITTI-13,由于摄像头移动会造成目标漂移,导致同一个目标的交叠率较低,若只采用交叠率作为关联依据易出现关联错误,本文采用自适应尺度颜色直方图,通过外观相似性提供较为准确的关联依据。表 3 是本文方法和其他方法在 2D MOT 2015 数据集上的实验结果对比。

表 3 2D MOT 2015 数据集上的对比结果

Table 3 Comparison results on 2D MOT 2015

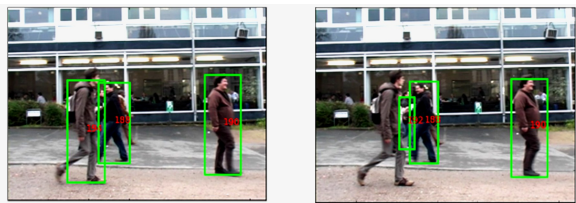
Methods	MOTA ↑	MT(%) ↑	FP ↓	FN ↓	ID sw ↓
RMOT ^[11]	18.6	5.3	12473	36835	684
TDAM ^[12]	33.0	13.3	10064	30617	464
MDP ^[8]	30.3	13.0	9717	32422	680
SORT ^[9]	33.4	11.7	7318	32615	1001
Proposed	34.1	12.8	7394	32361	729

图 6 和图 7 给出了采用 SORT 方法和本文方法在 2D MOT 2015 数据集上分别进行实验的部分实验效果图。对比两个方法的跟踪效果图可以发现,在出现目标漂移、目标交错以及目标间遮挡的情况时,本文提出的方法有更好的跟踪效果。在建立跟踪目标和检测目标间的一对一关系时,相似度融合的方式提高了关联过程的鲁棒性和准确性,提出的遮挡预测方法解决了短时目标间的遮挡问题,两种方法的结合使用在整体上提升了跟踪效果。



(a)TUD-Campus 第 8 帧

(b)TUD-Campus 第 9 帧

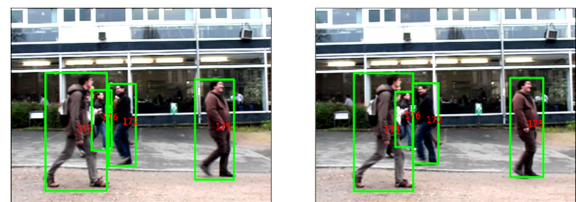


(c)TUD-Campus 第 10 帧

(d)TUD-Campus 第 11 帧

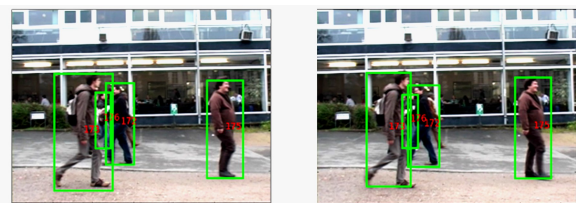
图 6 SORT 方法的多目标跟踪效果

Fig. 6 Results of MOT using SORT



(a)TUD-Campus 第 8 帧

(b)TUD-Campus 第 9 帧



(c)TUD-Campus 第 10 帧

(d)TUD-Campus 第 11 帧

图 7 本文方法的多目标跟踪结果

Fig. 7 Results of MOT using proposed method

结束语 本文采用基于卡尔曼滤波的多目标跟踪方法和高效的匈牙利方法建立跟踪目标和检测目标间的一对一关系。为了提高关联准确性,将融合交叠率和自适应尺度空间颜色直方图模型作为目标关联的依据,同时针对由于目标间的遮挡造成漏检从而导致跟踪失败的情况,提出了遮挡预测方法。实验结果表明,本文所提方法减少了由于目标交错、漂移以及目标间的遮挡导致的跟踪失败的情况,提高了跟踪准确性,在 2D MOT 2015 数据集上的平均准确率达到 34.1%,较已有多目标跟踪方法提升了跟踪效果。

(下转第 276 页)

- [3] CHAN T F, SHEN J. Mathematical Models for Local Nontexture Inpaintings [J]. *Siam Journal on Applied Mathematics*, 2002, 62(3): 1019-1043.
- [4] VESE L A, OSHER S J. Modeling Textures with Total Variation Minimization and Oscillating Patterns in Image Processing [J]. *Journal of Scientific Computing*, 2003, 19(1): 553-572.
- [5] GROSSAUER H. A Combined PDE and Texture Synthesis Approach to Inpainting [C] // *Computer Vision-ECCV 2004, European Conference on Computer Vision, Prague, Czech Republic, 2004*. DBLP, 2004: 214-224.
- [6] JI H L, YANG Q W. Image inpainting algorithm based on group-structured sparse representation [J]. *Computer Engineering and Applications*, 2016, 52(18): 14-17. (in Chinese)
计宏磊, 杨清文. 基于群结构稀疏表示的图像修复 [J]. *计算机工程与应用*, 2016, 52(18): 14-17.
- [7] AHARON M, ELAD M, BRUCKSTEIN A. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation [J]. *IEEE Transactions on Signal Processing*, 2006, 54(11): 4311-4322.
- [8] DONG W, SHI G, LI X. Nonlocal image restoration with bilateral variance estimation; a low-rank approach [J]. *IEEE Transactions on Image Processing*, 2013, 22(2): 700-711.
- [9] DONG W, ZHANG L, SHI G M, et al. Nonlocally centralized sparse representation for image restoration [J]. *IEEE Transactions on Image Processing*, 2013, 22(4): 1620-1630.
- [10] ZHANG J, ZHAO D, GAO W. Group-based sparse representation for image restoration [J]. *IEEE Transactions on Image Processing*, 2014, 23(8): 3336-3351.
- [11] XU Z, SUN J. Image inpainting by patch propagation using patch sparsity [J]. *IEEE Transactions on Image Processing*, 2010, 19(5): 1153-1165.
- [12] ZHANG L, ZHANG D, MOU X Q, et al. FSIM: A Feature Similarity Index for Image Quality Assessment [J]. *IEEE Transactions on Image Processing*, 2011, 20(8): 2378-2386.

(上接第 257 页)

参考文献

- [1] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards real time object detection with region proposal networks [C] // *International Conference on Neural Information Processing Systems*. MIT Press, 2015: 91-99.
- [2] HE K, ZHANG X, REN S, et al. Deep Residual Learning for Image Recognition [C] // *Computer Vision and Pattern Recognition*. IEEE, 2016: 770-778.
- [3] HOFMANN M, WOLF D, RIGOLL G. Hypergraphs for Joint Multi-view Reconstruction and Multi-object Tracking [C] // *Computer Vision and Pattern Recognition*. IEEE, 2013: 3650-3657.
- [4] DICLE C, CAMPS O I, SZNAIER M. The Way They Move: Tracking Multiple Targets with Similar Appearance [C] // *IEEE International Conference on Computer Vision*. IEEE, 2014: 2304-2311.
- [5] KIM C, LI F, CIPTADI A, et al. Multiple Hypothesis Tracking Revisited [C] // *IEEE International Conference on Computer Vision*. IEEE, 2015: 4696-4704.
- [6] REZATOFIGHI S H, MILAN A, ZHANG Z, et al. Joint Probabilistic Data Association Revisited [C] // *IEEE International Conference on Computer Vision*. IEEE, 2015: 3047-3055.
- [7] BAE S H, YOON K J. Robust Online Multi-object Tracking Based on Tracklet Confidence and Online Discriminative Appearance Learning [C] // *IEEE Conference on Computer Vision and Pattern Recognition*. IEEE Computer Society, 2014: 1218-1225.
- [8] XIANG Y, ALAHI A, SAVARESE S. Learning to Track; Online Multi-object Tracking by Decision Making [C] // *IEEE International Conference on Computer Vision*. IEEE, 2015: 4705-4713.
- [9] BEWLEY A, GE Z, OTT L, et al. Simple online and realtime tracking [C] // *IEEE International Conference on Image Processing*. IEEE, 2016: 3464-3468.
- [10] CHEN J G, QIN X S, MA L L. Fast GM-PHD Filter for Multi-target Tracking [J]. *Computer Science*, 2016, 43(3): 317-321. (in Chinese)
陈金广, 秦晓姍, 马丽丽. 快速多目标跟踪 GM-PHD 滤波算法 [J]. *计算机科学*, 2016, 43(3): 317-321.
- [11] JU H Y, YANG M H, LIM J, et al. Bayesian Multi-object Tracking Using Motion Context from Multiple Objects [C] // *Applications of Computer Vision*. IEEE, 2015: 33-40.
- [12] YANG M, JIA Y. Temporal Dynamic Appearance Modeling for Online Multi-Person Tracking [J]. *Computer Vision & Image Understanding*, 2015, 31(9): 626-633.
- [13] BENFOLD B, REID I. Stable multi-target tracking in real-time surveillance video [C] // *Computer Vision and Pattern Recognition*. IEEE, 2011: 3457-3464.
- [14] KUHN H W. The Hungarian method for the assignment problem [J]. *Naval Research Logistics*, 2005, 52(1): 7-21.
- [15] BREITENSTEIN M D, REICHLIN F, LEIBE B, et al. Online Multiperson Tracking-by-Detection from a Single, Uncalibrated Camera [J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2011, 33(9): 1820-1833.
- [16] YUAN D L, JI Q G. Multiple Object Tracking Algorithm via Collaborative Motion Status Estimation [J]. *Computer Science*, 2017, 44(Z11): 154-159. (in Chinese)
袁大龙, 纪庆革. 协同运动状态估计的多目标跟踪算法 [J]. *计算机科学*, 2017, 44(Z11): 154-159.
- [17] BIRCHFIELD S T, RANGARAJAN S. Spatiograms versus Histograms for Region-Based Tracking [C] // *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE, 2005: 1158-1163.
- [18] CONAIRE C O, O'CONNOR N E, SMEATON A F. An Improved Spatiogram Similarity Measure for Robust Object Localisation [J]. *IEEE International Conference on Acoustics*, 2007, 1(4): 1069-1072.
- [19] YANG B, NEVATIA R. Online Learned Discriminative Part-Based Appearance Models for Multi-human Tracking [C] // *European Conference on Computer Vision*. Springer-Verlag, 2012: 484-498.
- [20] LEMENAGER E, BOUET T, BRAIBANT V. Kalman filtering; a new approach for building global approximations. Application to the inverse optimization of an explosively formed penetrator (EEP) [J]. *Structural Optimization*, 1997, 14(2/3): 158-164.