

DOI: 10.7652/xjtuxb201810018

融合历史轨迹的智能汽车城市复杂环境 多目标检测与跟踪算法

隗寒冰¹, 陈尧¹, 贾志杰¹, 赖锋²

(1. 重庆交通大学车辆与汽车工程学院, 400074, 重庆; 2. 东风汽车集团有限公司技术中心, 430030, 武汉)

摘要: 针对现有智能汽车环境感知算法多根据特定类型目标设计,在处理目标遮挡、光照突变等城市复杂场景时识别准确率较低的问题,提出一种基于网状分类器与融合历史轨迹的多目标检测与跟踪算法。该算法考虑各目标之间的遮挡关系,利用具有目标融合功能的网状分类器对多尺度滑动窗获取的待检窗口进行多目标检测;历史检测结果基于目标特征关联通过计算目标长短轨迹和历史轨迹可靠性验证生成历史轨迹库,该轨迹库用于预测或融合新的检测结果;利用该检测跟踪结果更新网状分类器中的标准差分类器、最近邻分类器和历史轨迹信息,直至完成多目标长时跟踪。实验结果表明,本文算法在目标遮挡、光照变化和阴雨天气的复杂城市环境下均可实现多目标长时间检测跟踪,与 KITTI 数据集样本相比,平均准确率在 77.17%~81.32%之间,单帧图像平均耗时 0.05 s,具有较好的实时应用前景。

关键词: 智能汽车;城市环境;网状分类器;多目标检测;历史轨迹

中图分类号: TP273 **文献标志码:** A **文章编号:** 0253-987X(2018)10-0132-09

A Multi-Target Detection and Tracking Algorithm Incorporating Historical Trajectories for Intelligent Vehicles in Urban Complicated Conditions

WEI Hanbing¹, CHEN Yao¹, JIA Zhijie¹, LAI Feng²

(1. School of Mechatronics and Vehicle Engineering, Chongqing Jiaotong University, Chongqing 400074, China;

2. Dongfeng Motor Corporation Technology Center, Wuhan 430030, China)

Abstract: A new type of multi-target detection and tracking algorithm based on net classifiers and historical trajectories is proposed to focus the problem that most existing popular environmental perception algorithms for intelligent vehicles are designed for specific type of targets and they have been proved to be low recognition accuracy in complicated urban conditions where occlusion of different targets and abrupt light change are frequent. Images are scanned from a multi-scale sliding window and then fused by a fusion model according to occlusion relations among pedestrians, vehicles and traffic signs. Historical detection results are obtained based on correlation of target features through short trajectory generation, long trajectory generation and reliability verification of trajectories, and are used to generate a library of historical trajectories, which is used to predict and to fuse detection results in the next step. Then, the standard deviation classifier, the nearest neighbor classifier and historical trajectories are updated by new detection and tracking results. Multi-target long-time tracking in complicated conditions is

收稿日期: 2018-01-17。 作者简介: 隗寒冰(1979—),男,副教授。 基金项目: 国家自然科学基金资助项目(51305472); 重庆市自然科学基金资助项目(cstc2014jcyjA6005)。

网络出版时间: 2018-08-16 网络出版地址: <http://kns.cnki.net/kcms/detail/61.1069.T.20180816.1705.012.html>

realized in this way. Field experiment shows that the proposed algorithm exhibits excellent performance of multi-target long-time tracking in complicated urban conditions such as rainy, shade, luminance changes. The average recognition rates for KITTI database samples is up to 77.17%~81.32% and calculation time of single frame picture is only 0.05 s. These results show a promising algorithm for real-time application.

Keywords: intelligent vehicle; urban condition; net classifier; multi-target detection; historical trajectories

作为智能汽车环境感知的关键技术,行人、车辆和交通标识等目标检测算法已受到广泛关注,在此基础上开发的车道偏离、前撞预警等驾驶员辅助系统也已得到应用,但在道路拥堵、目标遮挡、光照突变和阴雨天气等复杂城市环境应用时仍面临严峻挑战。

较早的行人、车辆和交通标识检测方法均建立在目标轮廓特征基础上,如 Gavrila 等将行人外形、尺寸和对称性等轮廓信息作为先验知识,筛选出行人待选区域后对各待选区域进行识别^[1-2];Srinivasa 等利用 sobel 算子检测出横、纵向边缘后根据尺寸等先验条件进行过滤得到候选边缘,然后经分级过滤、直线型过滤得到最终边缘^[3];Bertozzi、Broggi 等在此基础上增加了边缘对称性判别^[4-5];Estable 等利用颜色、形状、直方图特征对候选区域交通标识进行检测^[6],此类方法虽检测效率高,但受光照、阴影、遮挡等影响明显;Bertozzi、Ouyang 采用基于特征模板配准方法,同样受遮挡限制^[7-8];也有研究采用基于目标特征和分类器以解决光照等环境影响^[9-10],如 Gabor、边缘对称性、轮廓等不同特征与 SVM、adaboost 等分类器结合,但复杂环境下的算法准确度急剧下降。

近年来,基于深度学习的目标识别算法逐渐成为研究热点^[11]。如 Girshick 采用自适应 ROI 区域和 R-CNN 深度网络进行图像特征提取^[12],并采用 SVM 分类器进行分类,不足之处是每个候选区域都需要前向卷积计算,无法满足实时性要求;虽然 Girshick 随后对此进行了改进,但全连接层制约了图像尺度缩放^[13];为同时满足尺度缩放及实时性要求,Ren 等利用全卷积神经网络(RPN)获取图像候选区域,并利用离线交替训练得到 Fast-R-CNN 网络识别目标^[14];Rdemon 等设计的 YOLO 算法遵循了 Google Net 的设计思想,将 RPN 和检测网络融合以实现端到端的目标识别^[15-16],该网络离线训练速度较快,但在线检测实时性较差;Liu、Janathan 等对网络设计细节进行了优化并提出 SSD 方法,将 YO-

LO 的方法推广到多尺度^[17-18];为保证信息的完整性,Janathan、Paszke、Zhao 等提出了基于反向卷积的上采样方法以保护图像边界^[18-20];Zhao 还提出了一种实时性强的语义分割网络(ICNet)^[21],但像素级语义分割的训练数据较少,语义边界难以准确获取、结果容易过拟合等问题仍制约着其实际应用。

为提高算法实时性,基于随机集概率密度模型^[22-23]、特征数据关联^[24-27]和运动信息配准的多目标跟踪算法相继出现,其中杨峰采用经典粒子滤波算法^[22],Okuma 等在此基础上对粒子的初始化分布和相似度判别进行了优化^[23],这一类方法在高速工况多目标长时跟踪具有较好效果,但对于城市工况效果不佳。Andriyenko、Mallick 采用了数据关联方法将多目标跟踪问题转化为全局最优问题,为提升数据关联性能引入轨迹约束^[24-25];Howard 等还采用了目标特征及其组成成分对关联数据结果进行优化^[26-27],这类方法虽然可实现道路拥挤时的多目标跟踪,但在多目标随机出现和消失时跟踪漏警率上升。融合跟踪、学习和检测的 TLD 算法^[28]将方差分类器、随机森林分类器和最近邻分类器串联,然后采用光流法进行目标跟踪,能有效解决单一目标长时间跟踪问题,但采用的随机森林分类器需要在线训练导致算法实时性低,计算耗时长达每帧 500 ms,特别是无法用于目标突然出现或者消失的场景。

本文针对道路拥挤、目标遮挡、光照突变等城市复杂环境下的多目标检测和跟踪问题,提出了基于网状分类器与融合历史轨迹的多目标检测与跟踪算法。受 TLD 在线学习反馈思想启发,利用在线学习的机制对分类器进行反馈训练,以满足长时间跟踪的周围环境变化需求。本文算法考虑各目标之间的遮挡关系,采用多尺度滑动窗口对待检图像进行扫描,网状分类器的检测结果通过目标融合器按照遮挡关系进行融合;其次根据历史检测结果,通过短轨迹生成、长轨迹生成和轨迹可靠性判别生成历史轨迹库,该历史轨迹库用于预测和融合新的检测结果;然后利用检测跟踪结果更新网状分类器中的标准差分

类器、最近邻分类器和历史轨迹库信息;最后迭代上述步骤直至实现复杂环境下的多目标长时跟踪。

1 基于网状分类器的多目标检测

车辆、行人和交通标识均可视为区域级刚性目标,可采用相同检测器与不同特征结合的方法实现多目标检测。本文以 Boost 分类器为基础,设计了一种网状多目标分类器,以实现对图像中行人、车辆和交通标识的检测。

1.1 图像特征选取

常用于目标检测的图像特征有 PCA 特征、LBP 特征、Haar 特征、HOG 特征、小波特征、shift 特征、纹理特征、不变矩特征等。HOG 特征可用来描述局部目标的表象和形状,适用于行人、车辆等轮廓复杂且具有一定规律的刚性目标。HOG 特征表示如下

$$\left. \begin{aligned} N(k) &= \sum_{x,y} \text{sgn}(\alpha_k(x,y),k) \\ C_x(k) &= \frac{\sum_{x,y} \text{sgn}(\alpha_k(x,y),k) \mathbf{G}(x,y)(x-x_c)}{\sum_{x,y} \text{sgn}(\alpha_k(x,y),k) \mathbf{G}(x,y)} \\ C_y(k) &= \frac{\sum_{x,y} \text{sgn}(\alpha_k(x,y),k) \mathbf{G}(x,y)(y-y_c)}{\sum_{x,y} \text{sgn}(\alpha_k(x,y),k) \mathbf{G}(x,y)} \end{aligned} \right\} \quad (1)$$

式中: $\alpha_k(x,y)$ 表示第 k 个方向对应的角度值; $N(k)$ 表示边缘梯度 $\mathbf{G}(x,y)$ 在第 k 个方向出现的次数总和; (x_c,y_c) 表示图像样本中心点坐标; $C_x(k)$ 、 $C_y(k)$ 分别为第 k 个方向上 $\mathbf{G}(x,y)$ 在 x 、 y 方向上分量的矢量和。

彩色直方图特征用于描述彩色图像色彩分布的统计规律,适用于交通标识等颜色差异明显的刚性目标。彩色直方图可分别用 R、G、B 这 3 个通道的直方图进行描述,其定义为

$$P(x_i) = \frac{N(x_i)}{\sum_{j=0}^{255} N(x_j)} \quad (2)$$

式中: x_i 、 x_j 分别为 R、G、B 通道下第 i 和 j 个像素值, $i,j \in (0,255)$; $P(x_i)$ 表示 R、G、B 通道下像素值 x_i 出现的概率; $N(\cdot)$ 表示 R、G、B 各通道下相应像素值之和。

1.2 特征概率密度优化的 Boost 分类器

行人、车辆、交通标识等刚性目标的正样本特征应集中在某一范围 ζ 内,可利用特征概率密度对 Boost 分类器初始权重进行优化。将正样本任意特

征列向量视为离散性随机变量,分布在 r 个区间上并服从概率分布,其概率分布函数为

$$P\{F \in \xi_r\} = \sum_{f_i \in \xi_r} P(F = f_i) \quad (3)$$

式中: P 为各区间 ζ_r 的概率。

选取样本分布概率 ω_r 对 Boost 分类器任意特征列的初始权值 $\mathbf{D}_n(f)$ 进行分配。对于任意特征列 j ,在第 n 次迭代时的弱分类器为

$$h_n(f_i, p_j, t_j) = \begin{cases} 1, & p_j \mathbf{F}_j(f_i) < p_j t_j \\ 0, & \text{其他} \end{cases} \quad (4)$$

式中: $\mathbf{F}_j(f_i)$ 是特征列 j 上第 f_i 个样本取值;偏置 p_j 取值为 1 或 -1 ; t_j 表示特征列 j 在第 n 次迭代时的训练阈值。

第 n 次迭代时最优偏置和阈值^[29]为

$$(p_j, t_j) = \operatorname{argmin} \left(\sum_{i=1}^m \mathbf{D}_n(f_i) (h_n(f_i, p, t) - y_i) \right) \quad (5)$$

式中: m 表示特征列 j 的样本数; y_i 表示第 i 个训练样本标签。

第 n 次迭代误差函数 R_n 可表示为

$$R_n = \sum_{i=1}^m \mathbf{D}_n(f_i) (h_n(f_i, p, t) - y_i) \quad (6)$$

经过多次迭代,更新样本权重使误差函数 R_n 取极小值

$$\mathbf{D}_{n+1}(f_i) = \frac{\mathbf{D}_n(f_i) e^{(-\frac{1}{2} \ln(\frac{1-R_n}{R_n}) y_i h_n(f_i, p, t))}}{Z_n} \quad (7)$$

式中: Z_n 为常数。

最终生成的强分类器 H_f 为

$$H_f = \operatorname{sgn} \left(\sum_{n=1}^T \frac{1}{2} \ln \left(\frac{1-\delta_n}{\delta_n} \right) h_n(f_i, p, t) \right) \quad (8)$$

式中: T 为最大迭代次数; δ_n 为分类器的分类误差。

式(8)得到的强分类器只能在准确率上满足目标检测需求,而误检率和实时性上表现较差。可将多个 Boost 分类器连接起来构成 Cascade 级联分类器,以解决单个 Boost 分类器误检率高、实时性差的问题。Cascade 分类器原理如图 1 所示。

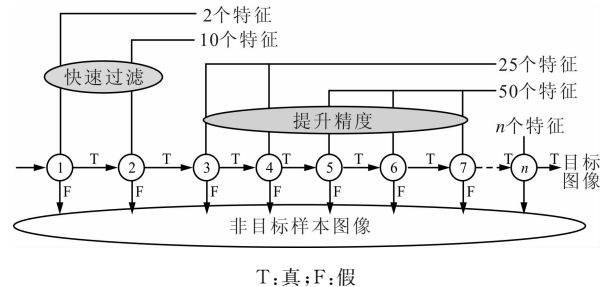


图 1 Cascade 分类器原理图

Cascade 分类器中每个待检图像只有通过当前层强分类器才可进入下层,位于 Cascade 起始位置的强分类器作用是快速过滤负样本,降低平均检测时间,中间层强和最后层分类器用于提升算法精度。如第 1 过滤器只过滤 2 个特征和 50% 非目标样本,共检测 99.8% 目标样本,第 2 过滤器只过滤 10 个特征和 50% 非目标样本,共检测 99.8% 目标样本,依次类推。Cascade 分类器训练过程中的正负样本比例取值在 1:4 到 1:6 之间。若比例过小,分类器会过度重视正样本,降低负样本权重,提高了分类器误警率;若比例过大,分类器会过度重视负样本,降低分类器准确率。

1.3 辅助分类器

将标准差分类器和最近邻分类器设置为辅助分类器。前者用于快速过滤图中负样本,后者用于判别检测结果是否为误检。标准差分类器可用待检窗口标准差与训练正样本集的标准差之间的差值表示。最近邻分类器用待检窗口特征矩阵 F 与正负样本集特征矩阵 S 的巴氏距离 s^+ 、 s^- 定义

$$f(s) = \begin{cases} 0, & \frac{s^+}{s^+ + s^-} < \omega \\ 1, & \frac{s^+}{s^+ + s^-} \geq \omega \end{cases} \quad (9)$$

$$s^+ = s^- = \left(\sum_{l=1}^n |F_{ij}^{(l)} - S_{ij}^{(l)}|^2 \right)^{\frac{1}{2}} \quad (10)$$

式中: ω 为设定阈值。

1.4 网状分类器

借鉴 Kalal 的不同类型检测器串联思想^[28],设计一种由标准差分类层、Cascade 分类层、最近邻分类层 3 层组成的网状分类器。该分类器主要针对行人、后端车辆、左侧车辆、右侧车辆、正方形交通标

识、圆形交通标识和三角形交通标识等 7 种典型情况进行多目标识别,网状分类器具体结构如图 2 所示。

合理假设图像中不可能在同一位置同时出现不同类型的交通标识或车辆,置信度判别模块只选取置信度最高的一个目标输出。考虑到道路拥挤时可能存在的多种目标遮挡现象,网状分类器利用遮挡关系判别器对多目标输出结果进行判别,然后输出检测结果。按照目标大小和出现位置,存在车辆遮挡车辆、行人遮挡车辆、交通标识遮挡车辆、交通标识遮挡行人和行人遮挡行人等 5 种遮挡关系。

1.5 多尺度滑动窗口检测

为提高算法实时性,根据目标图像的成像规律符合相机近大远小透视规律的特点,本文提出了一种多尺度滑动窗检测方法。该方法的基本思想是行人、车辆只能位于地面行驶,则运动轨迹成像规律可用过消失点的直线近似表示,位于空中的交通标识的成像规律可用直线近似表示。多尺度遍历扫描具体过程如下:①离线标记多组目标矩形框,记录各种目标在不同中心点下的尺寸大小,生成中心点与尺寸的历史数据图;②滑动窗口检测时,各待检测目标根据离线生成的历史数据图选取最佳尺寸窗口进行检测。扫描过程如图 3 所示,其中 L_1 、 L_2 、 L_3 和 L_4 表明目标运动轨迹。

2 基于历史轨迹关联的多目标跟踪

首先根据目标在图像中的位置、尺寸、置信度、类型和特征 5 个参数,建立相应的短轨迹相似度模型,并对多目标进行管理以将帧间目标连接成为短轨迹;然后根据短轨迹的运动特性,利用速度和预测目标重叠度建立相应的长轨迹相似度模型,并对各

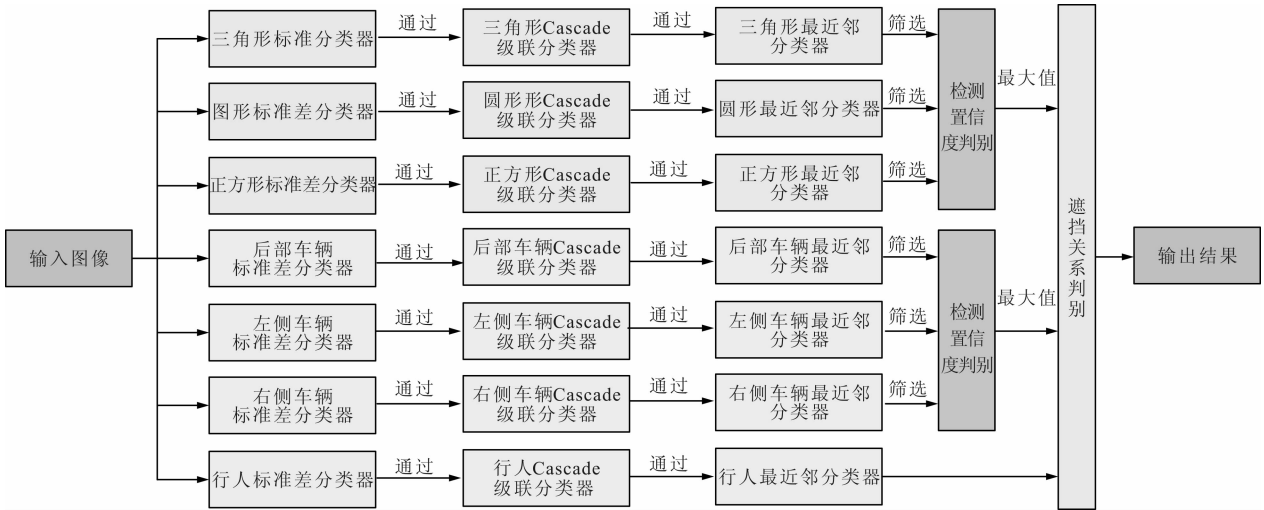


图 2 网状分类器原理示意图



图 3 多尺度滑动窗检测方式示意图

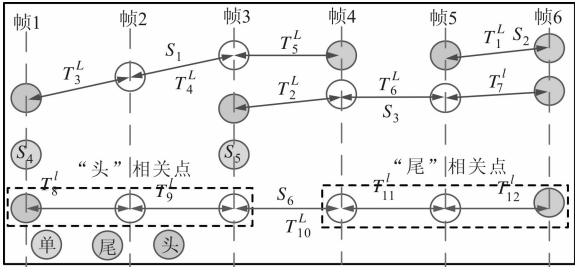


图 4 历史短轨迹生成原理

轨迹进行管理,以将短轨迹连接成为长轨迹;其次根据长轨迹插值求出丢失目标,生成历史轨迹信息库;最后将新检测目标与历史轨迹关联、历史轨迹预测、历史轨迹更新依次循环,最终实现多目标长时跟踪。

2.1 历史短轨迹生成

令 $z_i^t(u_i^t, v_i^t)$ 、 $s_i^t(w_i^t, h_i^t)$ 、 c_i^t 、 g_i^t 、 F_i^t 分别表示 t 时刻下第 i 个目标的位置坐标、尺寸大小、置信度、类型和特征矩阵。对于 t 时刻任意目标 i 与 $t + \Delta t$ 时刻任意目标 j 的相似度定义为

$$w_{ij} = P_z P_s P_c P_g P_f \tag{11}$$

式中: P_z 表示位置相似度,用中心点坐标距离定义; P_s 表示尺寸相似度,用目标尺寸定义; P_c 表示置信度相似度,用目标输出置信度定义; P_f 为特征矩阵相似度,用特征矩阵巴氏距离的指数定义; P_g 为类型相似度,相同类型取 1,不同类型则取 0。

$$P_z = \exp\left(-\frac{\sqrt{(u_i^t - u_j^{t+\Delta t})^2 + (v_i^t - v_j^{t+\Delta t})^2}}{\max(w_i^t, h_i^t) \Delta t}\right) \tag{12}$$

$$p_s = \exp\left(-\frac{\max((w_i^t - w_j^{t+\Delta t}), (h_i^t - h_j^{t+\Delta t}))}{\max(w_i^t, h_i^t) \Delta t}\right) \tag{13}$$

$$P_c = \exp\left(-\frac{c_i^t - c_j^{t+\Delta t}}{c_i^t \Delta t}\right) \tag{14}$$

$$P_f = \exp\left(-\sum_{x \in X} \sqrt{F_i^t \otimes F_j^{t+\Delta t}}\right) \tag{15}$$

根据式(11)计算结果,可得到协相似度矩阵对于 t 和 $t + \Delta t$ 时刻的所有目标进行管理。依次按照相似度大小对目标依次连接,直至符合连接阈值的目标均已连接完成,过程如图 4 所示,对应历史短轨迹生成结果见表 1。

令短轨迹的起点为“头”,终点为“尾”,“头”和“尾”附近的点称为“头、尾相关点”,只有一个点称为“孤立点”。同一轨迹起点的图像帧号不大于其终点的图像帧号,且起点和终点仅能连接 1 个目标。轨

迹连接过程总是从起点到终点,被连接短轨迹的起点必须大于连接轨迹的终点。

表 1 历史短轨迹生成结果

短轨迹序号	头帧数	尾帧数	短轨迹包含目标
S_1	1	4	T_3^L, T_4^L, T_5^L
S_2	5	6	T_1^L
S_3	3	6	T_2^L, T_6^L, T_7^L
S_4	1	1	S_4
S_5	3	3	S_5
S_6	1	6	$T_8^L \sim T_{12}^L$

2.2 历史长轨迹生成

参考目标点生成历史短轨迹的方法, Δt 时间内所有历史长轨迹中的任意两短轨迹之间的相似度 q_{ij} 可用正逆向预测速度差值、预测位置重叠度、预测目标特征矩阵相似度来定义

$$q_{ij} = P_v P_o P_f \tag{16}$$

如图 5 所示,令 v_i^s, v_i^e 表示第 i 个历史短轨迹头、尾处速度。根据第 i 个历史短轨迹尾处速度,前向预测出其在第 j 个历史短轨迹头处的速度 v_j^p 与位置及尺寸 d_j^p ,同理后向推测出 v_j^p, d_j^p 。根据 v_j^p 和第 j 个历史短轨迹头处真实速度、位置可计算出速度差 Δv_j^p 和前向预测重叠面积 ΔS_j^p ;根据 d_j^p 和第 i 个历史短轨迹尾真实速度、位置可计算出速度差 Δv_i^p 和后向预测重叠面积 ΔS_i^p 。历史短轨迹速度相似度可定义为

$$P_v = \exp\left(-\frac{\Delta v_i^p + \Delta v_j^p}{2\Delta\xi}\right) \tag{17}$$

式中: $\Delta\xi$ 表示历史短轨迹 i 的末端与历史短轨迹 j 的起始之间间隔帧数。预测位置重叠度相似度 P_o 。则定义为

$$P_o = \exp\left(-\frac{1}{2\Delta\xi}\left(\frac{\Delta S_i^p}{w_j^p h_j^p} + \frac{\Delta S_j^p}{w_i^p h_i^p}\right)\right) \tag{18}$$

P_f 定义同式(15)。

根据式(16)计算结果,可得到协相似度矩阵,并对 Δt 时间内所有轨迹进行管理。根据协相似度矩

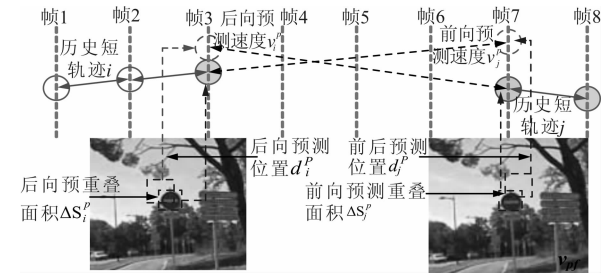


图 5 历史长轨迹生成原理

阵,依次按照相似度大小对轨迹及其可连接依次连接,直至符合连接阈值的轨迹均已连接完成,最后通过线性插值计算出丢失帧目标。最终生成的历史长轨迹如图 6 所示,每一条历史长轨迹均对应于相机坐标系下的长时跟踪目标。如图 6 中 1、3、5 箭头所指轨迹为场景中交通标志牌目标跟踪轨迹,2、4、6 箭头所指轨迹为车辆目标跟踪轨迹。

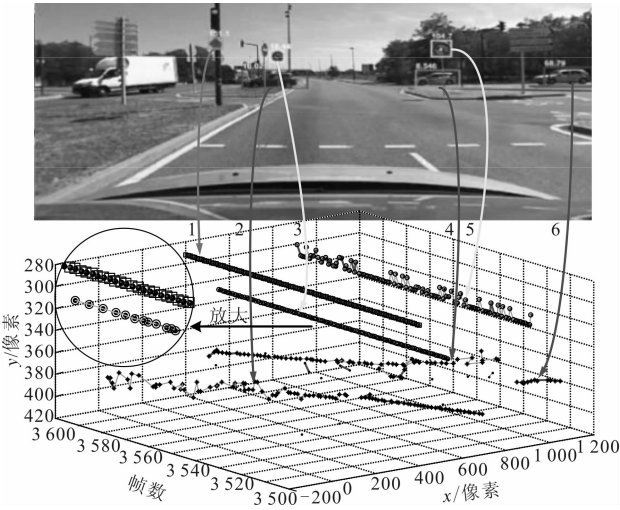


图 6 历史长轨迹与目标的对应关系

2.3 历史长轨迹可靠性验证

为提高长轨迹生成质量,应对生成的长轨迹进行可靠性验证。其可靠度用历史长轨迹长度 l 、检测比例 d 、轨迹目标置信度 c 三者定义

$$C = C_l C_d C_c \tag{19}$$

式中:长度置信度 C_l 用二值函数定义,当历史长轨迹长度 l 小于设定阈值时 C_l 取 0,否则取 1;检测比例置信度 C_d 用历史长轨迹中检测目标与插值目标的比值定义;检测置信度用轨迹中最大几个目标置信度 c 的均值确定,若轨迹可靠度小于设定阈值,则被判定为噪声。

2.4 长时跟踪与信息更新

将生成的历史轨迹与在线更新相结合,通过反复迭代检测、预测、学习、更新过程实现多目标的长时跟踪,其过程如图 7 所示。

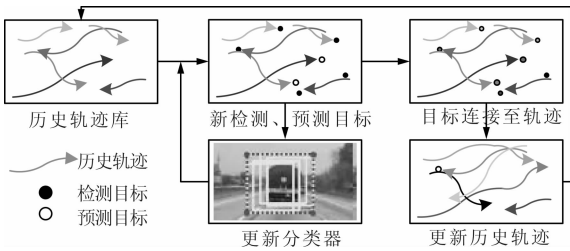


图 7 历史轨迹下的长时跟踪原理

算法对前 γ 帧图像只做检测计算,按 2.1~2.3 节所述生成历史轨迹。从第 $\gamma+1$ 帧起,执行检测、预测、学习、更新运算。第 $\gamma+1$ 帧中,若检测器检测结果可与历史轨迹连接,则将检测结果融入长轨迹,同时对标准差分类器和最近邻分类器的域值和权重值系数进行学习和更新;若结果无法融入任何轨迹,则被视为孤立点。连接完成后,若第 $\gamma+1$ 帧仍有检测结果未能成功连接到历史轨迹,则通过该检测结果的当前运动轨迹信息预测目标在下一帧中所在的位置。轨迹不能无限延长,每隔 $\Delta\gamma$ 帧历史轨迹库进行一次更新,生成最近 $\Delta\gamma$ 帧的历史轨迹。如此迭代计算,直至实现多目标长时跟踪。

3 道路实验结果与分析

3.1 城市环境下多目标检测跟踪结果

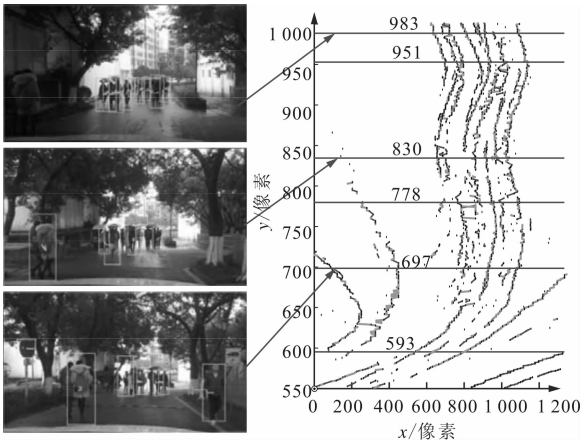
为验证算法准确性,选取涵盖直行、十字路口转向、校园内部道路等多种工况的山地城市环境进行实验。实验器材选用普通 USB 相机,分辨率为 1 280 像素×720 像素,图像采样频率 25 Hz,相机安装车内高度 1.35 m。车载工控机处理器为 i5-4210 M,主频 2.60 GHz,内存 8 GB。

图 8 所示为树荫和立交桥造成光照突变的城市道路多目标检测跟踪结果。图 8 中交通标识、车辆和行人检测与跟踪结果均用矩形框表示。

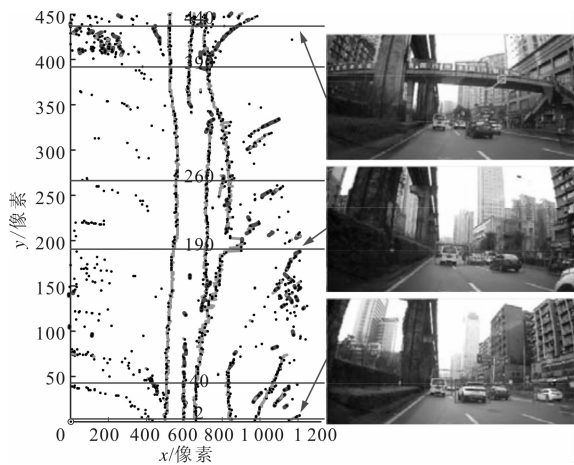
图 8a 中历史轨迹段对应的第 593、697、778 和 830 帧均为雨天城市路段密集人群下的多目标检测和跟踪结果。图 8b 中历史轨迹段对应的第 2、40、190 和 390 帧均为城市路段极其拥堵情况下的多目标检测和跟踪结果。可以看出,在存在遮挡的情况下,算法仍可有效实现多目标的检测和长时跟踪,最长跟踪时长可达到 2 min。

3.2 KITTI 数据集不同算法的对比结果

为对算法性能进行客观评价,利用公开数据库 KITTI 提供的不同算法目标检测结果进行对比,见表 2。可以看出,基于单一分类器的传统检测算法如 LSVM 等的准确率均在 50%~58% 范围,误检



(a)行人、标志牌多目标检测跟踪结果



(b)车辆目标检测跟踪结果

图 8 典型城市道路多目标检测跟踪结果

率较高,其原因是对于城市道路中较多出现的车辆、行人目标检测,基于特征识别和单一分类器的传统方法容易受到环境因素干扰。在深度学习网络基础

上开发的 iDST-VC、RV-CNN 等算法虽然准确率均超过了 90%,但这些算法为了提高准确性均采用了较深的网络结构,对硬件资源的要求苛刻,在使用 GPU 计算的前提下计算速度仅为 4 s/帧,显然无法应用于智能车实际道路环境感知。本文算法的平均准确率在 77.17%~81.32%之间,虽然低于深度学习算法,但相对于 LSVM 等传统算法有明显提升。更为重要的是,该算法在未采用 GPU 硬件加速下的平均单帧时间仅为 0.05 s,说明该算法在实时性和准确性之间取得较好的平衡,能够满足车载嵌入式处理器上的实时应用要求。

表 2 不同算法的性能对比结果

算法	准确率/%	单帧时间/s	硬件资源
iDST-VC	90.55	4.00	GPU
BM-NET	90.48	4.00	GPU
SAIT	90.36	0.15	GPU
RV-CNN	90.14	3.50	GPU
CNN	87.42	1.00	单核
spLBP	77.39	1.50	8 核
多目标检测跟踪	77.24	0.05	8 核
YOLOv2	69.01	0.03	GPU
LSVM-MDPM-us	56.10	10.00	4 核
Faster RCNN	56.58	0.11	GPU

图 9 给出了本文算法与同样采用传统特征识别和分类器的 LSVM-MDPM-us 算法的比较结果。场景涵盖了城市工况经常遇到的车辆和行人密集、目标交错重叠、阴影干扰和光线突变等情况。图 9a 为各种光照突变和成像距离变化等条件下的算法对比结果,图 9b 为场景中出现多个重叠行人目标的检



(a)光照突变、成像距离变化场景

(b)行人重叠遮挡场景

图 9 两种算法利用 KITTI 标准数据集检测结果的对比

测结果,图中黑色矩形框表示 LSVM-MDPM-us 算法检测结果,白色矩形框表示多目标检测跟踪算法结果。可以看出,在各场景下本文算法均可有效实现多目标长时检测跟踪,而 LSVM-MDPM-us 算法明显准确率较低,均未检测或跟踪到多个目标。

4 结 论

目标重叠、阴影干扰和光线突变等复杂城市工况下,智能车辆多目标检测和跟踪准确率下降,是制约其实际应用的最大挑战。受 TLD 多目标检测跟踪算法启发,本文提出了一种基于网状分类器与融合历史轨迹的多目标检测与跟踪算法。根据网状分类器检测结果,通过短轨迹生成、长轨迹生成和轨迹可靠性判别生成历史轨迹库;然后利用检测跟踪结果对网状分类器中的标准差分类器、最近邻分类器和历史轨迹信息进行更新,直至实现多目标长时跟踪。实车实验结果表明,该算法在目标遮挡、光照突变等复杂城市环境下均可实现多目标长时间检测跟踪。KITTI 样本平均准确率在 77.17%~81.32% 之间,单帧图像平均耗时 0.05 s,具有较好的实时应用前景。

参考文献:

- [1] GAVRILA D. Pedestrian detection from a moving vehicle [C]//Proceedings of the European Conference on Computer Vision. Berlin, Germany: Springer, 2000: 37-49.
- [2] GAVRILA D M, GIEBEL J. Shape-based pedestrian detection and tracking [C]//Proceedings of the Intelligent Vehicle Symposium. Piscataway, NJ, USA: IEEE, 2002: 8-14.
- [3] SRINIVASA N. Vision-based vehicle detection and tracking method for forward collision warning in automobiles [C]//Proceedings of the Intelligent Vehicle Symposium. Piscataway, NJ, USA: IEEE, 2002: 626-631.
- [4] BERTOZZI M, BROGGI A, FASCIOLI A, et al. Stereo vision-based vehicle detection [C]//Proceedings of the Intelligent Vehicles Symposium. Piscataway, NJ, USA: IEEE, 2000: 39-44.
- [5] BROGGI A, CERRI P, ANTOELLO P C. Multi-resolution vehicle detection using artificial vision [C]//Proceedings of the Intelligent Vehicles Symposium. Piscataway, NJ, USA: IEEE, 2004: 310-314.
- [6] ESTABLE S, SCHICK J, STEIN F, et al. A real-time traffic sign recognition system [C]//Proceedings of the Intelligent Vehicles Symposium. Piscataway, NJ, USA: IEEE, 1994: 213-218.
- [7] BERTOZZI M, BROGGI A, CASTELLUCCIO S. A real-time oriented system for vehicle detection [J]. Journal of Systems Architecture, 1997, 43(1): 317-325.
- [8] OUYANG W, WANG X. Joint deep learning for pedestrian detection [C]//Proceedings of the International Conference on Computer Vision. Piscataway, NJ, USA: IEEE, 2013: 2056-2063.
- [9] BETKE M, HARITAOGLU E, DAVIS L S. Multiple vehicle detection and tracking in hard real-time [C]//Proceedings of the Intelligent Vehicles Symposium. Piscataway, NJ, USA: IEEE, 1996: 351-356.
- [10] PAPAGEORGIOU C P, OREN M, POGGIO T. A general framework for object detection [C]//Proceedings of the International Conference on Computer Vision. Piscataway, NJ, USA: IEEE, 1998: 555-562.
- [11] STALLKAMP J, SCHLIPSING M, SALMEN J, et al. Man vs. computer: benchmarking machine learning algorithms for traffic sign recognition [J]. Neural Networks the Official Journal of the International Neural Network Society, 2012, 32(2): 323-332.
- [12] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2014: 580-587.
- [13] GIRSHICK R. Fast R-CNN [C]//Proceedings of the International Conference on Computer Vision. Piscataway, NJ, USA: IEEE, 2015: 1440-1448.
- [14] REN Shaoqing, HE Kaiming, ROSS G, et al. Faster R-CNN: towards real-time object detection with region proposal networks [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2015, 39(6): 1137-1149.
- [15] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [C]//Proceedings of the Computer Vision and Pattern Recognition. Berlin, Germany: Springer, 2015: 779-788.
- [16] REDMON J, FARHADI A. YOLO9000: better, faster, stronger [C]//Proceedings of the Computer Vision and Pattern Recognition. Berlin, Germany: Springer, 2017: 21-25.
- [17] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector [C]//Proceedings of the

- European Conference on Computer Vision. Berlin, Germany: Springer, 2016: 21-37.
- [18] JONATHAN L, EVAN S, TREVOR D. Fully convolutional networks for semantic segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 4(39): 640-651.
- [19] PASZKE A, CHAURASIA A, KIM S, et al. ENet: a deep neural network architecture for real-time semantic segmentation [C]// Proceedings of the Computer Vision and Pattern Recognition. Berlin, Germany: Springer, 2017: 21-25.
- [20] ZHAO Hengshuang, SHI Jianping, QI Xiaojuan, et al. Pyramid scene parsing network [C]// Proceedings of the Computer Vision and Pattern Recognition. Berlin, Germany: Springer, 2017: 21-25.
- [21] ZHAO Hengshuang, QI Xiaojuan, SHEN Xiaoyong, et al. ICNet for real-time semantic segmentation on high-resolution images [C]// Proceedings of the Computer Vision and Pattern Recognition. Berlin, Germany: Springer, 2017: 21-25.
- [22] 杨峰,王永齐,梁彦,等. 基于概率假设密度滤波方法的多目标跟踪技术综述 [J]. 自动化学报, 2013, 39(11): 1944-1956.
- YANG Feng, WANG Yongqi, LIANG Yan, et al. A survey of PHD filter based multi-target tracking [J]. ACTA Automatic Sinica, 2013, 39(11): 1944-1956.
- [23] OKUMA K, TALEGHANI A, FREITAS N D, et al. A boosted particle filter: multitarget detection and tracking [C]// Proceedings of the European Conference on Computer Vision. Berlin, Germany: Springer, 2004: 28-39.
- [24] ANDRIYENKO A, SCHINDLER K. Globally optimal multi-target tracking on a hexagonal lattice [C]// Proceedings of the European Conference on Computer Vision. Berlin, Germany: Springer, 2010: 466-479.
- [25] MALLICK M, KRANT J, SHALON J. Multi-sensor multi-target tracking with out-of-sequence measurements [C]// Proceedings of the 6th International Conference of Information Fusion. Piscataway, NJ, USA: IEEE, 2003: 672-679.
- [26] HOWARD A, PADGETT C, LIEBE C C. A multi-stage neural network for automatic target detection [C]// Proceedings of the IEEE World Congress on Computational Intelligence. Piscataway, NJ, USA: IEEE, 1998: 4-9.
- [27] GRINBERG M, OHR F, BEYERER J. Feature-based probabilistic data association (FBPDA) for visual multi-target detection and tracking under occlusions and split and merge effects [C]// Proceedings of the International Conference on Intelligent Transportation Systems. Piscataway, NJ, USA: IEEE, 2009: 1-8.
- [28] KALAL Z, MIKOLAJCZYK K, MATAS J. Tracking-learning-detection [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2011, 34(7): 1409-1422.
- [29] HU Weiming, HU We, MAYBANK S. AdaBoost-based algorithm for network intrusion detection [J]. IEEE Transactions on Systems Man and Cybernetics: Part B Cybernetics, 2008, 38(2): 577-583.

[本刊相关文献链接]

- 刘毅,秦贵和,赵睿,等. 车载控制器局域网络安全协议. 2018,52(5):94-100. [doi:10.7652/xjtuxb201805014]
- 邓涛,罗俊林,韩海硕,等. 混合动力汽车工况识别自适应能量管理策略. 2018,52(1):77-83. [doi:10.7652/xjtuxb201801012]
- 张亮修,王宇,吴光强,等. 汽车阻尼可调半主动悬架混杂模型预测控制. 2017,51(11):56-164 [doi:10.7652/xjtuxb201711022]
- 李超超,向建华,王慧敏. 汽车保险杠系统吸能盒结构参数对低速碰撞下吸能特性的影响. 2017,51(10):77-81. [doi:10.7652/xjtuxb201710013]
- 章军辉,李庆,陈大鹏,等. 基于 BP 神经网络的纵向避撞安全辅助算法. 2017,51(7):140-147. [doi:10.7652/xjtuxb201707020]
- 严珍,王斌,徐俊,等. 双参数组合优化的复合电源模式切换控制策略. 2016,50(11):129-135. [doi:10.7652/xjtuxb201611020]
- 卢礼华,陆建辉,刘志峰,等. 汽车坐盆安全气囊对假人伤害的仿真及优化. 2016,50(9):146-152. [doi:10.7652/xjtuxb201609023]
- 许广灿,徐俊,李士盈. 电动汽车振动能量回收悬架及其特性优化. 2016,50(8):90-95. [doi:10.7652/xjtuxb201608015]
- 卢礼华,刘志峰,陆建辉. 帘式安全气囊有限元建模及侧碰撞系统仿真. 2016,50(7):104-109. [doi:10.7652/xjtuxb201607016]
- 王斌,徐俊,曹秉刚,等. 电动汽车的多模式复合电源能量管理自适应优化. 2015,49(12):130-136. [doi:10.7652/xjtuxb201512021]
- 邓涛,林椿松,李亚南,等. 采用 NSGA-II 算法的混合动力能量管理控制多目标优化方法 [J]. 西安交通大学学报, 2015, 49(10):143-150. [doi:10.7652/xjtuxb201510023]
- 王斌,徐俊,曹秉刚,等. 采用模拟退火算法的电动汽车复合电源能量管理系统优化. 2009,43(6):10-14. [doi:10.7652/xjtuxb201508015]

(编辑 武红江)