

doi: 10.19562/j.chinasae.qcgc.2024.12.012

# 车载雷达关键场景的数据采集与标记方法\*

黄楷博<sup>1</sup>, 邓伟文<sup>1</sup>, 王莹<sup>2</sup>, 赵蕊<sup>1</sup>, 丁娟<sup>3</sup>(1. 北京航空航天大学交通科学与工程学院, 北京 100191; 2. 吉林大学计算机科学与技术学院, 长春 130015;  
3. 嘉兴南湖学院, 嘉兴 314000)

**[摘要]** 车载雷达的虚警和漏报是影响自动驾驶系统安全可靠的关键因素之一, 因此需要大量带标签的测试数据开展针对性的研究。但虚警和漏报的发生概率较低, 且雷达目标状态不稳定导致雷达目标难以标记。对此, 本文首先根据雷达虚警和漏报的产生机制设计能高效获取雷达关键数据的测试方案。然后通过构建关联度函数, 以量化雷达目标与场景目标之间的关联并使用遗传算法优化该函数, 在此基础上建立雷达目标的自动标记方法。最后通过实采数据验证本文方法的有效性。实验结果表明所提出的方法能高效获取关键的虚警和漏报数据, 本文的标记方法也能准确识别出场景目标对应的雷达目标, 并区分出虚警目标和真实目标。

**关键词:** 自动驾驶; 毫米波雷达; 虚警; 漏报; 目标标记

## Data Collection and Annotation Method for Radar on Some Key Scenarios

Huang Kaibo<sup>1</sup>, Deng Weiwen<sup>1</sup>, Wang Ying<sup>2</sup>, Zhao Rui<sup>1</sup> & Ding Juan<sup>3</sup>1. School of Transportation Science and Engineering, Beihang University, Beijing 100191;  
2. College of Computer Science and Technology, Jilin University, Changchun 130015;  
3. Jiaxing Nanhu University, Jiaxing 314000

**[Abstract]** False alarm and missed alarm of automotive radar are key factors affecting the safety and reliability of autonomous driving systems, thus requiring a large amount of labeled test data for targeted research. However, the occurrence probability of false alarm and missed alarm is low, and the unstable status of radar targets makes it difficult to label them. Therefore, in this paper, firstly efficient test schemes are designed to obtain key radar data based on the generation mechanism of radar false alarm and missed alarm. Then, by constructing a correlation function to quantify the correlation between radar targets and scene targets and using genetic algorithms to optimize this function, an automatic labeling method for radar targets is established. Finally, the effectiveness of the proposed method is verified through real data acquisition. The experimental results show that the proposed method can efficiently obtain crucial false alarm and missed alarm data. The labeling method in this paper can accurately identify radar targets corresponding to scene targets and distinguish between false alarm and real targets.

**Keywords:** autonomous driving; millimeter-wave radar; false alarm; missed alarm; object labeling

## 前言

自动驾驶可以减少交通事故, 提升交通效率, 是未来交通的发展方向<sup>[1]</sup>。毫米波雷达具有探测距离远和全天候工作的特点, 是自动驾驶中不可或缺的

传感器<sup>[2]</sup>。雷达检测结果中存在遗漏目标或识别出虚假目标的情况, 即漏报和虚警。虽然虚警和漏报发生的概率较低, 但这种非理想的感知结果会导致雷达输出的目标数量和类别不稳定, 增加感知结果的复杂度, 是影响自动驾驶系统可靠性的主要原因。因此, 包含虚警和漏报的场景成为雷达的目标识别、

\* 浙江省“领雁”研发攻关计划项目(2023C01238)和浙江省“尖兵”研发攻关计划项目(2023C01133)资助。

原稿收到日期为 2024 年 05 月 11 日, 修改稿收到日期为 2024 年 06 月 13 日。

通信作者: 王莹, 副教授, 博士, E-mail: wangying\_jlu@163.com; 丁娟, 副教授, 博士, E-mail: juan.ding@panosim.com。

跟踪等感知算法和雷达建模等技术关注的关键场景<sup>[3]</sup>。

基于监督学习的数据驱动方法是感知算法和雷达建模等应用的主要实现方法<sup>[4-5]</sup>。数据驱动方法的效果受数据质量的影响非常大<sup>[6-7]</sup>,高质量的数据集应包含均衡的特征分布和准确的标签。其中不均衡的数据集会导致使用该数据集建立的模型偏向多数的特征、忽视少数特征,标签的错误或者缺失会导致监督学习模型向错误的目标优化,在实际应用中导致重大错误。对于应用于雷达目标识别、跟踪和雷达建模的雷达数据集,应该包含充足的虚警和漏报场景的数据和准确的场景真值的标签和雷达目标的标签。雷达目标的标签包括雷达目标与场景目标的对应关系以及虚警目标和真实目标的标签,如图1所示。

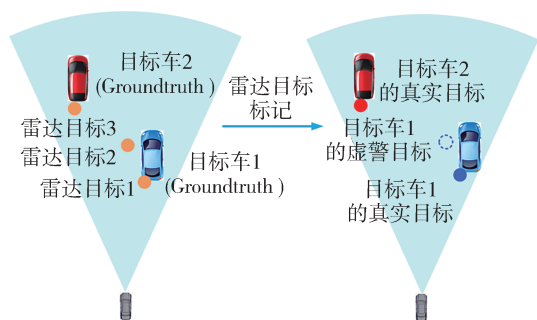


图1 雷达数据集的标签

目前关于雷达数据集的研究有很多,牛津大学、利物浦大学和韩国KAIST大学分别建立了包括多种天气和交通场景的开源数据集 Oxford Radar RobotCar<sup>[8]</sup>、RaDiCaL<sup>[9]</sup>和MulRan<sup>[10]</sup>。华盛顿大学、Motional公司、奥尔登堡大学和法国的研究者同样采集了多种场景下的雷达和相机等传感器数据构建了CRUW<sup>[11]</sup>、nuScenes<sup>[12]</sup>、RadarScenes<sup>[13]</sup>和CARRADA<sup>[14]</sup>等数据集,并建立了行人、自行车、汽车、货车等多种目标的标签。但当前的雷达数据集中只包含常见驾驶场景的雷达数据,没有针对虚警和漏报场景的测试方案。这导致数据集中的虚警和漏报场景的数据非常少,不能满足高性能的雷达算法和雷达模型开发的需求。另一方面,虽然部分数据集建立了雷达数据的标签,但是其标签中没有区分雷达的真实目标和虚警目标,且普遍没有场景目标的真实参数<sup>[15]</sup>。这些都极大地限制了这些数据的实用价值。

为解决上述的情况,本文首先根据雷达探测机

理分析虚警漏报的产生机制,明确自动驾驶场景中发生虚警和漏报的条件并在此基础上制定能高效获取关键数据的采集方案。然后在实车场景中基于雷达和惯导传感器采集数据并通过坐标变换、时间同步等预处理等方法,获取与雷达输出数据同步的场景目标真实参数。随后提出一种雷达输出目标的标记方法,该方法根据雷达机理建立函数量化雷达目标与场景目标的关联度,并使用遗传算法优化其中的量化参数,在此基础上建立雷达的标记方法识别出场景目标对应的雷达目标并分辨出虚警目标和真实目标。最后通过实采数据验证本文的采集方法和标记方法的有效性。

## 1 雷达关键数据采集方案

### 1.1 雷达的虚警和漏报产生机制

雷达向外界发射信号并接收由目标反射回雷达的回波信号。回波信号强度 $P_r$ 由式(1)计算。其中, $P_0$ 为天线的发射功率, $G_t$ 为发射天线的方向性系数, $G_r$ 为接收天线的方向性系数,是与目标角度相关的函数,代表雷达对不同方向的信号强度的增益, $R$ 为雷达与目标的距离, $\lambda$ 为雷达工作频率对应的波长, $\sigma$ 为目标的雷达散射截面(radar cross-section, RCS),反映目标对电磁波的反射能力。

$$P_r = \frac{\sigma P_0 G_t G_r \lambda^2}{(4\pi)^3 R^4} \quad (1)$$

目标的回波信号与噪声等干扰信号混合到一起被雷达接收,经由雷达信号处理模块进行目标检测。正常情况下目标的回波信号强度大于干扰信号强度,雷达正确地检测出目标<sup>[16]</sup>。

#### 1.1.1 漏报的原理分析

当回波信号强度过小时雷达无法从混合信号中分辨出目标回波信号,导致出现漏报。根据式(1),目标的距离 $R$ 越大,信号衰减越大。目标的角度越大天线的增益 $G_t$ 、 $G_r$ 越小。除距离和角度外,目标的RCS越小,反射信号的强度越低。RCS由目标的材质、外形等参数决定。其中目标的外形特指目标正面朝向雷达部位的外轮廓,由目标自身的三维外形、目标的姿态决定。雷达能有效检出常见目标的角度和距离范围称为雷达的有效探测范围。但自动驾驶场景中目标面向雷达的外轮廓会因为目标间的遮挡而发生巨大的变化。由此可见,目标超出有效探测范围,或目标之间存在遮挡时是导致漏报的主要原

因,如图2所示。

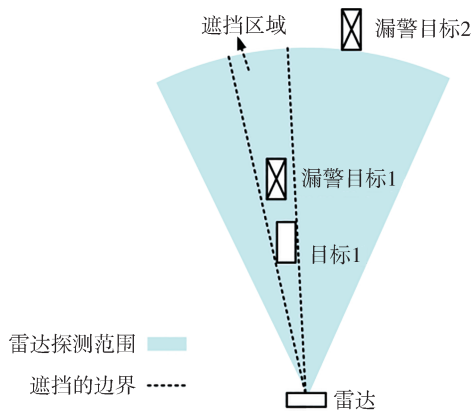


图2 雷达目标漏报的场景

### 1.1.2 虚警的原理分析

雷达接收到的回波信号中包含干扰信号,若干干扰信号的强度超过雷达目标检测的阈值,雷达会将干扰信号识别为虚警目标。高强度干扰信号通常由多径引起,如图3所示。自动驾驶场景中目标不规则的三维外形将入射的毫米波信号向多个方向反射,反射信号入射到其他目标外表面或者地面等地方后发生二次或多次反射,多次反射后的部分信号被雷达接收。但由式(1)可知,当多径信号入射到的目标的RCS较大、多径路程较短即目标之间距离较近时,多径产生的衰减很小,可以产生达到雷达的检测阈值的高强度干扰信号。此外,目标的速度虽然不直接影响虚警目标的回波信号强度,但雷达的目标检测算法会参考雷达回波信号中的速度信息来分辨不同的目标,所以速度参数也会影响虚警发生的概率。由此可见,当场景目标之间距离接近、RCS较大以及存在速度差时,容易产生虚警。

### 1.2 雷达关键数据采集方案

建立高质量的雷达数据集应充分考虑场景因素

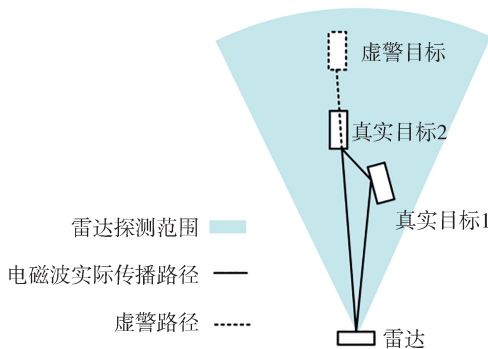


图3 雷达目标虚警的场景

对雷达探测结果的影响,增加存在虚警和漏报的场景占比以保证数据集中样本的均衡性,并覆盖自动驾驶中的常见工况以增加数据的实用价值。

首先根据虚警和漏报的产生机制明确容易产生虚警和漏报的场景。测试中导致虚警和漏报的原因主要是多径和遮挡。符合要求的驾驶行为包括车辆之间的跟随、变道、超车、并行行驶和对向行驶等,如图4所示。其中车辆之间的跟随、变道和超车的行为伴随着车辆之间相对位置的变化,容易产生遮挡而发生漏报。并行、对向行驶、超车等驾驶行为中目标之间的距离接近,产生的多径信号强度衰减减小,车辆之间还存在速度差,容易产生虚警。此外,目标车辆与雷达的相距较远的场景容易因回波信号衰减而导致漏报,也是重要的测试场景。

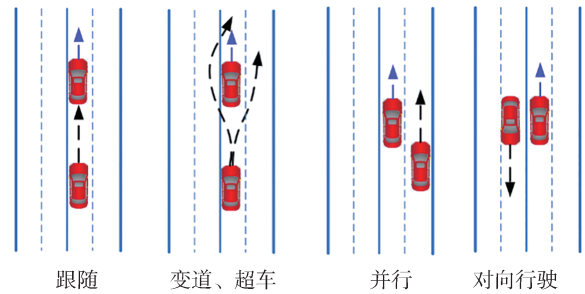


图4 车辆间的交互

自动驾驶的常见工况主要为高速公路场景、郊区道路场景、城市道路场景和低速场景。常见工况中的目标主要为车辆,常见工况下车辆相对雷达的行驶速度和距离如表1所示。

表1 常见工况下车辆相对雷达的距离和速度

场景	高速场景	郊区场景	城市场景	低速场景
速度/(km·h <sup>-1</sup> )	60-120	40-80	30-60	0-20
距离/m	33.3-66.6	33.3-66.6	16.6-83.5	1-20

参考虚警和漏报的产生场景和实际工况的雷达数据采集方案如图5和表2所示。图中在实际工况的速度和距离上将车辆的速度参数设置为低速、中速和高速3种模式。并考虑到雷达的探测范围对漏报的影响,将车辆的距离设置为近距离、中等距离和远距离3种模式。车辆之间的距离分为近距离和中远距离2种模式,这里取2个车道的距离7m作为分割的阈值。该数据采集方案可以在覆盖常见场景的同时获取足够的包含虚警和漏报的雷达数据。

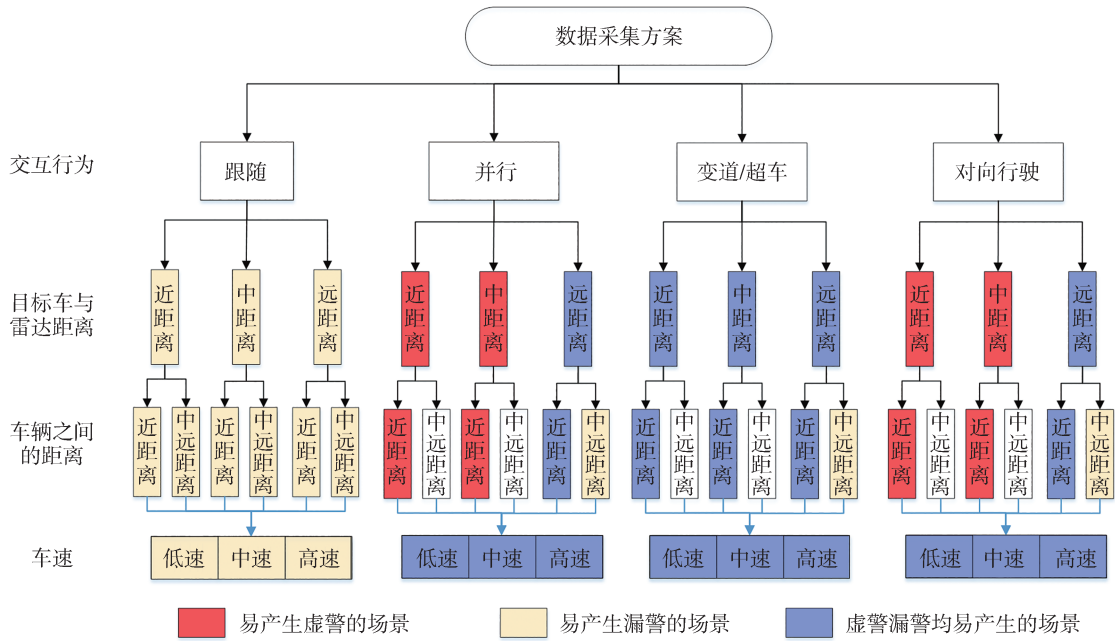


图5 雷达数据采集方案

表2 车辆的运动参数

速度/(km·h <sup>-1</sup> )	低速	中速	高速
	0-30	30-60	60-120
距离/m	近距离	中距离	远距离
	0-20	20-83.5	83.5-160
车辆之间的距离/m	近距离	中远距离	
	1-7	7-160	

### 1.3 数据采集方法

高质量的雷达数据包含场景目标的真值数据和雷达输出的目标及目标参数。为获取准确的场景真值与雷达数据,在测试车的正前方安装雷达,并连接控制器局域网(controller area network, CAN)分析仪和上位机以获取雷达数据,如图6所示。在目标车辆和测试车辆上安装包含惯导设备并连接上位机获取高精度的车辆的运动数据。在此基础上,通过车辆间的相位运动状态获得目标车相对于雷达的运动参数。

### 1.4 数据预处理

#### 1.4.1 坐标变换

雷达输出如图7所示的雷达坐标系下的目标的参数( $RCS, R_x, R_y, v_x, v_y$ )。 $RCS$ 表示雷达目标的能量强度, $R_x, R_y$ 分别为横向距离和纵向距离, $v_x, v_y$ 分别为横向速度和纵向速度。惯导设备输出地球经纬度坐标系下,定位天线所在位置的经纬度坐标、东向速度、北向速度和车辆的航向角信息等车辆的运动信

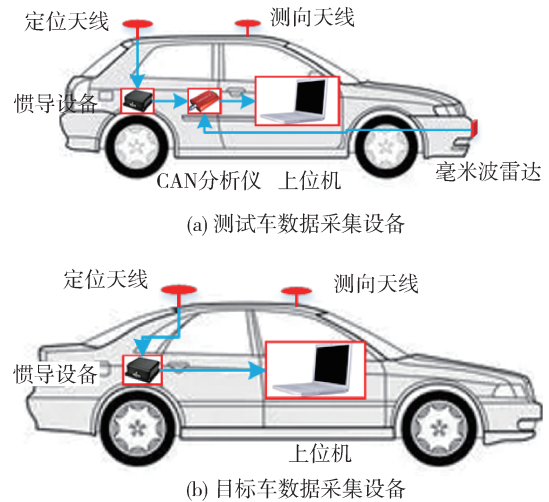


图6 雷达与场景数据采集设备

息。为获取雷达坐标系下的场景真值,由车辆间惯导的相对运动数据获取目标车相对雷达的运动参数,并将运动参数由地球坐标系下变换到雷达坐标下。

(1)根据测试车上定位天线与雷达的相对位置获得雷达在地球经纬度坐标系下的位置、距离和航向角参数。

(2)根据雷达和目标车惯导的位置、速度和航向角信息计算地球坐标系下目标车相对于雷达的速度、距离和航向角。

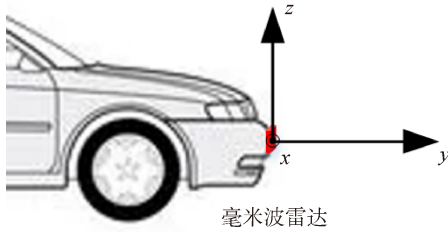


图7 雷达坐标系

(3)将地球坐标系下的场景目标参数变换到雷达坐标系下,建立与雷达输出参数坐标统一的场景目标速度、位置和航向角信息。

#### 1.4.2 数据同步

实际工作中,惯导以 10 ms 的周期输出包含全球卫星定位系统(GPS)时间的数据,由此可以同步不同车辆之间的惯导数据,但雷达以 50–100 ms 的周期输出不包含时间戳的数据。为实现雷达数据和惯导数据的同步,将测试车惯导与雷达的输出数据用同一根 CAN 总线输出到上位机进行存储,雷达和惯导输出数据会按照接收的时间顺序保存到上位机。根据接收的数据顺序和惯导数据的时间戳对雷达数据的探测时刻进行近似计算,获取雷达近似的 GPS 时间。最后用插值的方法计算雷达探测时刻对应的惯导真值数据。

如图 8 所示,雷达探测一次的周期内,惯导已完成多次探测。每条雷达数据前后均有惯导数据,设这两条惯导数据的 GPS 时间戳分别为  $GPStime_1$  和

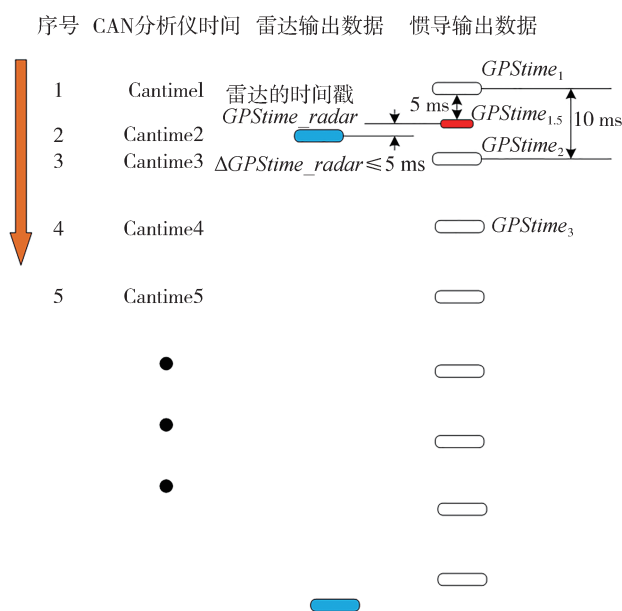


图8 雷达数据与惯导数据同步

$GPStime_2$ ,该条雷达数据的接收时刻  $GPStime\_radar$  位于前后 2 个惯导接收时刻的区间内  $[GPStime_1, GPStime_2]$ ,惯导 2 个探测时刻的中间时刻为  $GPStime_{1.5}$ ,取  $GPStime\_radar$  等于  $GPStime_{1.5}$ ,可以保证误差在 5 ms 内。

$$GPStime_{1.5} = (GPStime_1 + GPStime_2) / 2 \quad (2)$$

$$|GPStime\_radar - GPStime_{1.5}| \leq 5 \text{ ms} \quad (3)$$

最后对  $GPStime_1$  和  $GPStime_2$  时刻的测试车和目标车惯导数据进行插值,获得  $GPStime\_radar$  时刻的惯导输出数据,实现雷达数据与惯导数据的同步。测试误差方面,测试方案中目标的最大速度为 33.3 m/s,在目标车按照该最大速度行驶时,5 ms 产生的最大测距误差为 0.17 m,属于可接受的范围。

$$P_{IMU\_GPStime\_radar} = (P_{IMU\_GPStime1} + P_{IMU\_GPStime2}) / 2 \quad (4)$$

## 2 基于机理和遗传算法的目标标记方法

高质量雷达数据集包含雷达目标的准确标签。本文提出一种基于机理的目标标记方法,该方法可以在任意场景中标记场景目标对应的雷达输出的目标,并分辨雷达输出目标中的真实目标和虚警目标。

根据雷达探测机理可知,雷达输出的真实目标和虚警目标均由场景目标引起,每个场景目标对应的真实目标数量为 1 个或 0 个,对应的虚警目标可以为任意数量。其中真实目标由场景中的目标对雷达发射的毫米波信号的直接反射引起。真实目标的 RCS、距离和速度参数 ( $RCS_t, v_{xt}, v_{yt}, R_{xt}, R_{yt}$ ) 接近场景目标真值 ( $RCS_{gt}, v_{xgt}, v_{ygt}, R_{xgt}, R_{ygt}$ )。虚警目标由电磁波在场景中的目标上的二次或多次反射引起,其参数值 ( $RCS_g, v_{xg}, v_{yg}, R_{xg}, R_{yg}$ ) 也接近场景目标的参数值,但与场景目标参数之间的差距相对于真实目标较大。同样,雷达目标与对应的场景目标的参数差值应小于与其他场景目标的参数差值。本文根据场景目标与雷达目标的差值量化雷达目标与场景目标之间的关联度。根据关联度的差异匹配雷达目标与场景目标,并根据场景目标与不同雷达目标之间的关联度差异分辨真实目标与虚警目标。

### 2.1 建立关联度函数

#### 2.1.1 建立关联函数量化目标间关联

由于雷达输出中各参数的动态范围和误差有很大的差异,无法通过简单的差值比较量化雷达目标与场景目标的关联度。本文将雷达目标与场景目标

的差值加权,建立评价雷达目标与场景目标关联度的式(5)。其中 $f_{co}$ 的数值越小,代表关联度越大, $w_{i0}$ 、 $w_{i1}$ 、 $w_{i2}$ 、 $w_{i3}$ 、 $w_{i4}$ 为各个参数的权值,须选择合适的权值以保证计算结果的有效性。

$$f_{co} = w_{i0} \Delta RCS + w_{i1} \Delta v_x + w_{i2} \Delta v_y + w_{i3} \Delta R_x + w_{i4} \Delta R_y \quad (5)$$

雷达的测速和测距精度代表了其检测速度和距离参数的误差上限,即限定了真实目标与对应的场景目标的速度和距离参数的最大差值( $\Delta v_{xr \max}$ 、 $\Delta v_{yr \max}$ 、 $\Delta R_{xr \max}$ 、 $\Delta R_{yr \max}$ )。同时也在一定程度上反映了场景目标与雷达目标的参数差值所代表的关联度,由此可以获得精度与关联度函数权值的关系。

$$\begin{cases} w_{i1} \propto \frac{1}{\Delta v_{xr \max}}; w_{i2} \propto \frac{1}{\Delta v_{yr \max}} \\ w_{i3} \propto \frac{1}{\Delta R_{xr \max}}; w_{i4} \propto \frac{1}{\Delta R_{yr \max}} \end{cases} \quad (6)$$

### 2.1.2 基于遗传算法优化权值

可信的关联度函数需要准确的权值,本文使用遗传算法进一步优化参数权值。遗传算法是一种受自然遗传机制启发的优化算法,通过模拟自然选择、遗传和变异过程来解决复杂的搜索和优化问题<sup>[17]</sup>。通常包括以下步骤<sup>[18]</sup>。

- (1)初始化:随机生成一个初始种群。
- (2)适应度评价:评估每个个体的适应度。
- (3)选择:根据个体的适应度进行选择。
- (4)交叉:选择的个体通过交叉(重组)操作产生后代。

(5)变异:对新的染色体进行随机的基因变异,帮助算法跳出局部最优解。

(6)新一代种群形成:新生成的个体形成新一代种群,替代原有种群中的一些或全部个体。

(7)终止条件:算法重复上述过程,直到满足预定的终止条件。

遗传算法在多个领域显示出其强大的搜索能力,尤其是在多目标优化问题上,非常适合用于本文中的多权值求解<sup>[19]</sup>。为构建适用于本文权值求解的适应性函数,首先从雷达数据集中选出少量的典型场景的数据,人工标记出这些数据中场景目标对应的雷达目标,以及虚警目标和真实目标。假设一共选出 $L$ 帧数据,某帧数据中一共有 $M$ 个雷达目标,其中 $N$ 个是对应于不同场景目标的真实目标。权值优化问题的适应性函数的目标是保证每个真实目标与

对应的场景目标的差值加权 $f_{real\_n}$ 小于其他任一雷达目标与该场景目标的差值加权 $f_{obj\_m}$ 。则适应性函数的公式为对 $L$ 帧数据中,每个真实目标与对应场景目标的 $f_{real\_n,l}$ 与其他目标与场景目标的 $f_{obj\_m,l}$ 的差值,若差值小于0则代表满足要求,将差值置零,然后将所有结果求和。

$$Fitness(w) = \sum_{l=1}^L \left( \sum_{n=1}^N \left( \sum_{m=1}^{M-1} \max(0, f_{real\_n,l} - f_{obj\_m,l}) \right) \right) \quad (7)$$

为了提高优化效率,在遗传算法优化权值时,将式(6)中的雷达精度的倒数作为初始值。通过迭代求解优化权值,建立关联度计算函数。

### 2.2 雷达目标标记方法

对任意的交通场景,首先基于机理设定对应于场景目标的雷达目标阈值范围,然后在阈值范围内使用关联度函数选出匹配的真实雷达目标,在剩下的目标中,根据关联度值将虚警目标与场景目标进行匹配。标记按照图9中的步骤进行。

#### (1)设置真实目标的阈值范围

雷达真实目标与场景目标的速度和距离参数相差在( $\Delta v_{xr \max}$ 、 $\Delta v_{yr \max}$ 、 $\Delta R_{xr \max}$ 、 $\Delta R_{yr \max}$ )范围内,由此可以确立场景目标对应的真实目标的阈值范围:( $v_{xgt} \pm \Delta v_{xr \max}$ 、 $v_{ygt} \pm \Delta v_{yr \max}$ 、 $R_{xgt} \pm \Delta R_{xr \max}$ 、 $R_{ygt} \pm \Delta R_{yr \max}$ )。

#### (2)建立各场景目标的真实目标集

分别筛选出满足各个场景目标阈值的所有雷达目标,建立各个场景目标的真实目标集。不同场景目标的真实目标集可能有重叠的情况。

#### (3)判断场景目标是否漏报

各场景目标对应的真实目标只可能存在于真实目标集中,若某个场景目标的真实目标集为空集,则该场景目标没有对应的真实雷达目标,为漏报状态。

#### (4)标记场景目标对应的真实雷达目标

在各个场景目标的真实目标集中,使用关联度计算式(5)获取雷达目标与场景目标的关联度,其中关联度最大的雷达目标为该场景目标对应的真实目标。

#### (5)剔除真实目标集中重叠的目标

判断不同场景目标对应的真实雷达目标是否相同。对于多个场景目标对应同一个真实雷达目标的情况。根据该雷达目标与几个场景目标的关联度进行排序,关联度最大的场景目标与该雷达目标存在对应关系,并返回第(2)步,从其余场景目标的真实目标集中剔除掉该目标并重新进行以上过程。

### (6)建立虚警目标集

完成所有真实目标的匹配后,从雷达输出的目标列表中,剔除掉所有真实目标,建立一个全部为虚警目标的虚警目标集。

### (7)计算场景目标对应的虚警目标

按照各虚警目标与场景目标的关联度进行排序,将各个虚警目标分别匹配到关联度最大场景目标。完成所有目标的标记工作。

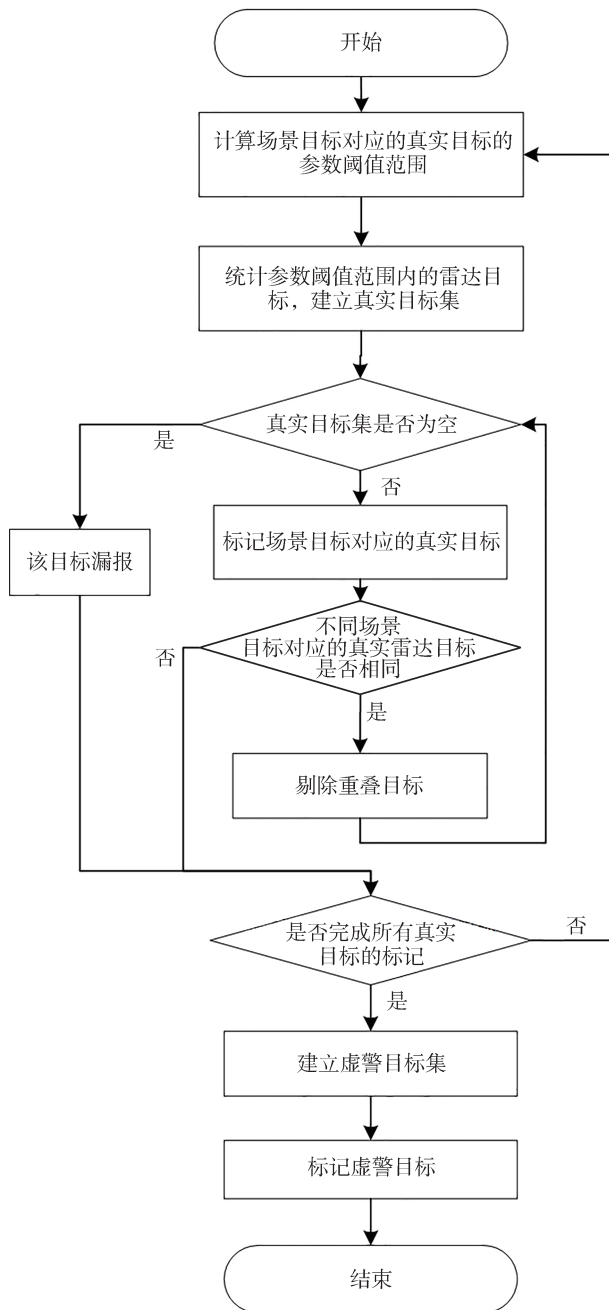


图9 目标标记流程图

## 3 结果验证

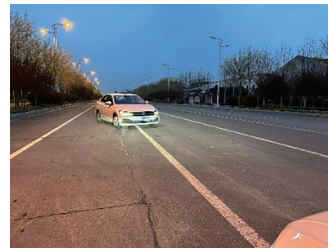
本文按照数据采集方案实采回波数据并对采集到的数据进行坐标变换和时间同步,对处理后的数据进行标记处理。验证本文的数据采集方案和雷达目标标记方法的有效性。

### 3.1 实车数据采集

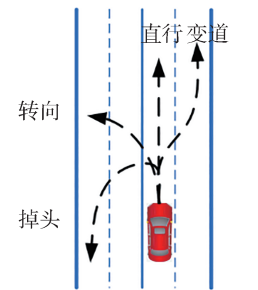
本文使用大众宝来作为测试车和目标车,测试车正前方安装大陆 ARS408 雷达,测试车和目标车车内安装华测 G610 惯导设备,车顶安装惯导的定位天线和测向天线,如图 10(a)所示。大陆 ARS408 雷达以 72 ms 的周期输出探测结果,惯导以 10 ms 的周期输出车辆运动信息,并将数据实时存储到上位机。



(a) 实车设备安装



(b) 数据采集场景



(c) 对照场景的车辆行驶轨迹

图10 数据采集

数据采集在中国江苏省南京市郊区进行,测试时使用2辆目标车在测试车前方按照数据采集方案中的轨迹和运动参数行驶,由雷达和惯导同时采集数据。将雷达和惯导数据进行时间同步和坐标变换,共获得雷达数据和场景目标真值数据 63 650 帧。与此同时,为对比验证数据采集方案的有效性设置了对照的采集方案,即不考虑车辆间的交互行为,两个目标车均按照图 10(c)所示正常的驾驶行为行驶,共采集数据 16 721 条。

### 3.2 数据标记与结果验证

通过数据处理和坐标变换获得场景目标的真实参数。在电磁仿真软件中构建车辆的三维结构,通过仿真获取车辆在典型角度下的RCS值用于构建关联度函数。

根据雷达的探测精度,通过式(6)获取本文的关联度函数式(5)的初始权值和真实目标的参数范围。由大陆ARS-408雷达的产品手册可以获得大陆ARS-408雷达的测速精度为0.028 m/s;测距精度为0.05-0.4 m;测角精度为0.1°-5°。测距和测角最大带来2.8 m的距离误差。实际应用中,还需要考虑车身的长度和车身的旋转等因素对雷达检测到的目标距离和速度的影响。这里考虑探测误差和目标车的车身长度带来的影响,设置真实目标的横向距离和纵向距离与车辆相差的最大阈值 $\Delta R_{xr \max}$ 和 $\Delta R_{yr \max}$ 为7.8 m。考虑到目标车上车身不同位置速度不一致的情况,这里设置横向速度和纵向速度的阈值 $\Delta v_{xr \max}$ 和 $\Delta v_{yr \max}$ 为0.25 m/s。

手动标记典型工况下的10帧数据,根据式(7)建立遗传算法的适应性函数,使用遗传算法优化关联度函数的权值。优化参数设置种群大小为200,遗传代数300,交叉概率0.7,变异概率0.08,迭代次数设置为300。优化获得的权值如表3所示。

表3 关联度函数权值

参数	$w_{i0}$	$w_{i1}$	$w_{i2}$	$w_{i3}$	$w_{i4}$
初始值	1	4	4	0.13	0.13
优化值	0.21	1.8	5.3	0.56	0.97

构建关联度函数后,根据标记方法中的步骤分别标记考虑交互行为的雷达数据集和不考虑交互行为的雷达数据集,结果如表4所示。可以看出,相对于不考虑交互行为的方案,根据虚警漏报的产生机理提出的车辆交互行为可以采集到大量的虚警和漏报数据,其中漏报的数据占比提升约5倍,虚警数据占比的提升超过2倍,证明该数据采集方案可以有效地获取虚警和漏报数据。

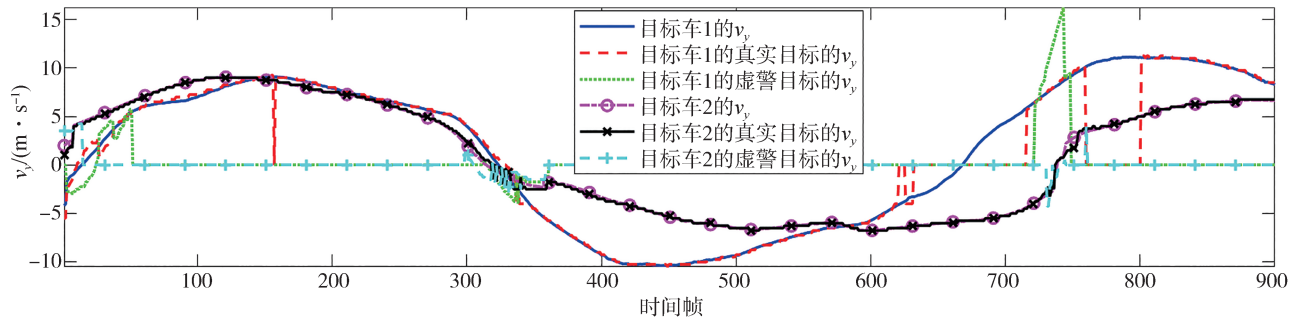
表4 雷达数据标记结果

数据集	场景目标	数据量	漏报数据数量	漏报概率/%	虚警数据数量	虚警概率/%
考虑车辆交互	目标车1	63 650	15 550	24.43	15 129	23.77
	目标车2		13 506	21.22	12 920	20.30
不考虑车辆交互	目标车1	16 721	847	5.07	1 560	9.33
	目标车2		602	3.60	1 393	8.33

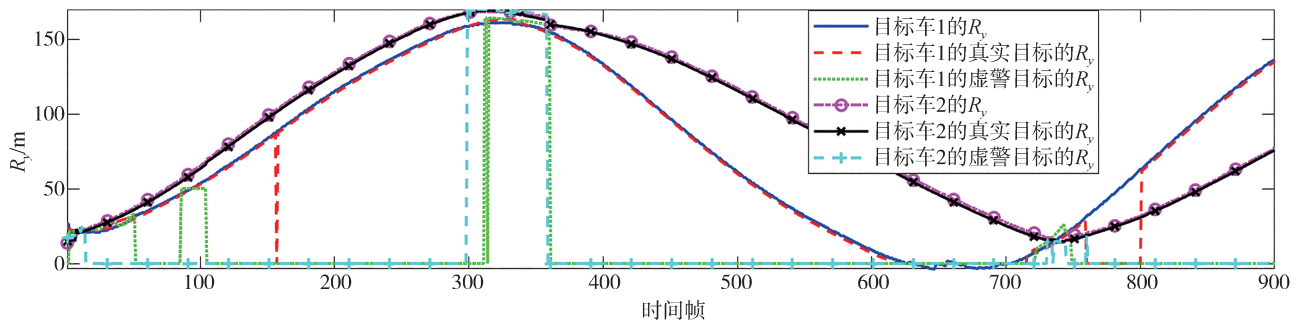
图11(a)和图11(b)中展示了2辆目标车和与其对应的真实目标和虚警目标的纵向距离值和纵向速度值。图中可以看出,目标车对应的真实目标的参数与场景真值非常接近,虚警目标的变化趋势与对应的目标车一致但其参数相对于真实目标有较大的误差,但两者都明显区别于另外一辆目标车的场景真值。表明该标记方法能有效地分辨出不同场景目标对应的雷达目标,且能区分出真实目标和虚警目标。此外,从 $R_r$ 图的1-100帧、300-360帧和720-760帧中可以清晰地看出,当2个目标车接近时会产生虚警目标。当目标车超出雷达的有效探测范围或者目标之间存在遮挡时会出现漏报的情况,如图中的620-800帧中第1辆目标车的真实目标展现出的结果。该结果也证明了数据采集方案的有效性。图11(c)和图11(d)中则分别展示了目标车1和目标车2与雷达目标的关联度值。可以看出,图中的结果与标记策略一致,目标车与其对应的真实目标的关联度值保持最小,目标车与对应的虚警目标的关联度小于与另一个目标车的虚警目标与真实目标。该结果证明了关联度函数的有效性。

## 4 结论

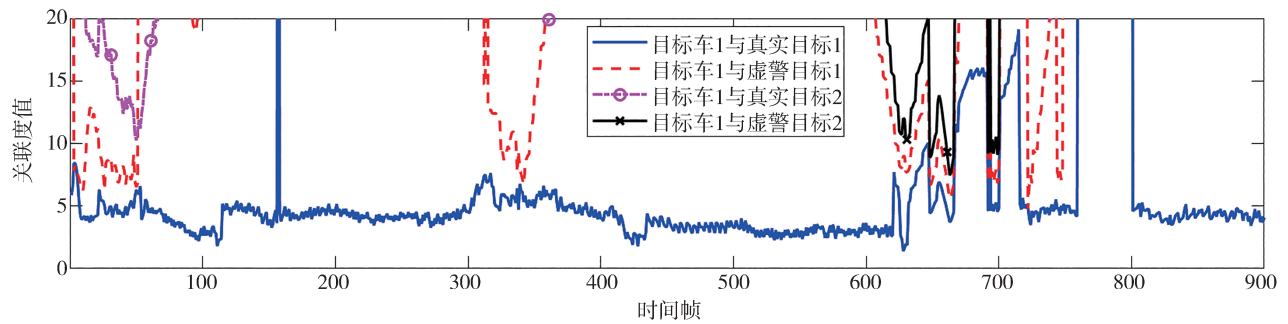
自动驾驶雷达传感器输出结果中的虚警和漏报现象是对自动驾驶可靠性的重大挑战,是感知算法和雷达建模的重点关注场景。当前感知算法和雷达建模越来越依赖数据驱动方法,但是当前的雷达数据集中缺少虚警和漏报的场景,且缺乏对数据的标记。针对这种情况,本文首先分析了雷达探测结果中虚警和漏报的产生机制,并将虚警漏报的产生机制与自动驾驶场景中的因素关联,设计了能高效获取虚警和漏报数据的雷达数据采集方案。随后提出了一种能同步获取雷达数据和场景中目标真值参数的数据采集和处理方法。最后基于机理与遗传算法量化雷达目标与场景目标的关联,从雷达目标中分辨出虚警目标与真实目标,并识别出雷达目标与场景目标的对应关系。文中实采雷达数据验证了本文的数据采集方案和标记方法的有效性。本文的方法可以用于具有复杂交通流的场景,对于雷达感知和建模有很大的帮助,在未来的研究中本研究团队会进一步关注本文的方法在实际的雷达感知和雷达建模上的应用。



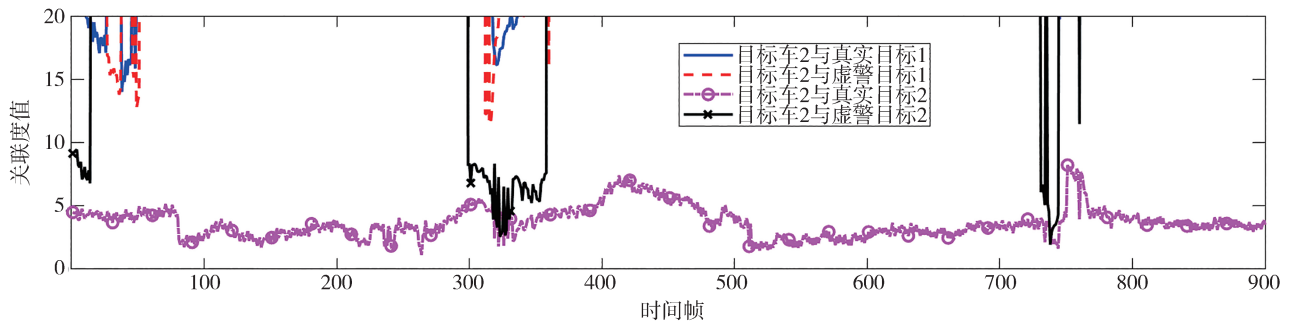
(a) 场景目标与雷达目标的纵向速度



(b) 场景目标与雷达目标的纵向距离



(c) 场景目标1与雷达目标的关联度变化曲线



(d) 场景目标2与雷达目标的关联度变化曲线

图11 标记结果示意图

参考文献

[1] FRIEDRICH B.The effect of autonomous vehicles on traffic[M]// MAURER M, GERDES J C, LENZ B, et al. Autonomous driv- ing: technical, legal and social aspects. Berlin, Heidelberg: Springer Berlin Heidelberg, 2016:317-334.

[2] MEINL F, STOLZ M, KUNERT M, et al.An experimental high performance radar system for highly automated driving [C].

- 2017 IEEE MTT-S International Conference on Microwaves for Intelligent Mobility (ICMIM). IEEE, 2017: 71-74.
- [3] GONG Y, XUE Y, DONG Q, et al. Quality evaluation of image dataset based on label file [C]. 2021 8th International Conference on Dependable Systems and Their Applications (DSA): 639-646.
- [4] MA Yifang, WANG Zhenyu, YANG Hong, et al. Artificial intelligence applications in the development of autonomous vehicles: a survey [J]. IEEE/CAA J. Autom. Sinica, 2020, 7(2): 315-329.
- [5] 詹军, 董学才, 洪峰, 等. 智能汽车传感器实时功能模型及验证 [J]. 汽车工程, 2019, 41(7): 731-737, 743.
- ZHAN J, DONG X C, HONG F, et al. Real-time functional model and verification of intelligent vehicle sensors [J]. Automotive Engineering, 2019, 41(7): 731-737, 743.
- [6] LI Y, CHAO X. Toward sustainability: trade-off between data quality and quantity in crop pest recognition [J]. Frontiers in Plant Science, 2021, 12.
- [7] LEDERER A, CAPONE A, UMLAUFT J, et al. How training data impacts performance in learning-based control [J]. IEEE Control Systems Letters, 2020, 5: 905-910.
- [8] BARNES D, GADD M, MURCUTT P, et al. The Oxford Radar RobotCar Dataset: a radar extension to the Oxford RobotCar Dataset [C]. 2020 IEEE International Conference on Robotics and Automation (ICRA): 6433-6438.
- [9] LIM T, MARKOWITZ S, DO M. RaDiCaL: a synchronized FMCW radar, depth, IMU and RGB camera data dataset with low-level FMCW radar signals [J]. IEEE Journal of Selected Topics in Signal Processing, 2021, 15: 941-953.
- [10] KIM G, PARK Y, CHO Y, et al. MulRan: multimodal range dataset for urban place recognition [C]. 2020 IEEE International Conference on Robotics and Automation (ICRA): 6246-6253.
- [11] WANG Y, JIANG Z, LI Y, et al. RODNet: a real-time radar object detection network cross-supervised by camera-radar fused object 3D localization [J]. IEEE Journal of Selected Topics in Signal Processing, 2021, 15: 954-967.
- [12] CAESAR H, BANKITI V, LANG A, et al. NuScenes: a multi-modal dataset for autonomous driving [C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR): 11618-11628.
- [13] SCHUMANN O, HAHN M, SCHEINER N, et al. RadarScenes: a real-world radar point cloud data set for automotive applications [C]. 2021 IEEE 24th International Conference on Information Fusion (FUSION): 1-8.
- [14] OUAKNINE A, NEWSON A, REBUT J, et al. CARRADA dataset: camera and automotive radar with range-angle-doppler annotations [C]. 2020 25th International Conference on Pattern Recognition (ICPR): 5068-5075.
- [15] 李勇滔, 孙晨旭, 郑伟光, 等. 基于毫米波雷达与视觉融合的碰撞预警 [J]. 汽车工程, 2023, 45(9): 1666-1676.
- LI Y T, SUN C X, ZHENG W G, et al. Collision warning based on fusion of millimeter wave radar and vision [J]. Automotive Engineering, 2023, 45(9): 1666-1676.
- [16] LI X, DENG W, ZHANG S, et al. Research on millimeter wave radar simulation model for intelligent vehicle [J]. Int. J. Automat. Technol., 2020, 21: 275-284.
- [17] MITCHELL M. An introduction to genetic algorithms [M]. Cambridge, Mass.: MIT Press, 1996.
- [18] WHITLEY L. A genetic algorithm tutorial [J]. Statistics and Computing, 199, 44: 65-85.
- [19] GOLDBERG D. Genetic algorithms in search optimization and machine learning [M/OL]. 1988. <https://doi.org/10.5860/choice.27-0936>.