

DOI: 10.19650/j.cnki.cjsi.J2311824

毫米波雷达微弱行人轨迹跟踪-预测一体化方法^{*}

方鑫¹,何敏¹,黄大荣²,张振源³,葛亮¹

(1. 西南石油大学机电工程学院 成都 610500; 2. 安徽大学人工智能学院 合肥 230601;
3. 重庆交通大学信息科学与工程学院 重庆 400074)

摘要:针对城市道路场景微弱行人目标雷达回波信号极易被强背景杂波淹没,导致目标轨迹跟踪及预测失效难题,提出了一种基于调频连续波-多输入多输出毫米波雷达的微弱行人轨迹跟踪-预测一体化方法。首先,利用递归贝叶斯检测前跟踪算法,直接从未经阈值处理的雷达三维原始频谱数据中提取目标运动轨迹,解决了传统阈值决策信息丢失带来的跟踪性能下降问题,并在此基础上提出了一种基于Transformer的端到端轨迹预测模型,进一步挖掘隐藏在跟踪轨迹中的时空相关性,完成了微弱行人目标轨迹的精准预测。实验结果表明,方法在信噪比大于-20 dB时,预测轨迹的平均位移误差和最终位移误差分别小于0.706、1.215 m,均优于高斯过程、长短期记忆网络等传统方法。

关键词: 行人轨迹预测;低信噪比;毫米波雷达;检测前跟踪

中图分类号: TN95 TH89 **文献标识码:** A **国家标准学科分类代码:** 510.40

Integrated trajectory tracking and prediction method for weak pedestrian with millimeter wave radar

Fang Xin¹, He Min¹, Huang Darong², Zhang Zhenyuan³, Ge Liang¹

(1. School of Electromechanical Engineering, Southwest Petroleum University, Chengdu 610500, China;
2. School of Artificial Intelligence, Anhui University, Hefei 230601, China; 3. School of Information
Science and Engineering, Chongqing Jiaotong University, Chongqing 400074, China)

Abstract: The radar echoes from weak pedestrians are easily submerged by strong background clutters in urban road traffic scenarios, resulting in the failure of trajectory tracking and prediction. To address this problem, this article proposes an integrated trajectory tracking and prediction method for weak pedestrians based on frequency-modulated continuous wave-multiple input multiple output millimeter wave radar. Firstly, the recursive Bayesian track-before-detect algorithm is utilized to directly extract target motion trajectory from the non-thresholding three-dimensional radar raw spectrum, which avoids the tracking performance degradation caused by the information loss from the traditional threshold-decision process. On this basis, this article proposes a Transformer-based end-to-end trajectory prediction model to further explore the spatiotemporal correlations of tracking trajectory and achieves accurate trajectory prediction of a weak pedestrian. Experimental results show that, when the signal-to-noise ratio is greater than -20 dB, the average displacement error and final displacement error of the predicted trajectory of the proposed method are less than 0.706 and 1.215 m, respectively, which are all superior to traditional methods such as Gaussian process and long short-term memory network.

Keywords: pedestrian trajectory prediction; low signal-to-noise ratio; millimeter wave radar; track-before-detect

0 引言

近年来,自动驾驶汽车的应用受到汽车行业越来越

多的关注。然而,行人作为交通场景中的主要参与者,对其未来运动轨迹的合理推理和预测对于自动驾驶和道路安全至关重要。因此,为了避免道路交通中的行人事故,有必要采用视觉^[1]、激光雷达^[2]、毫米波雷达^[3]先进的车

收稿日期:2023-08-23 Received Date: 2023-08-23

^{*} 基金项目:国家自然科学基金(62303386, 62273065, 62003064, 52174209)、重庆市自然科学基金(cstc2020jcyj-msxmX0797)、重庆市教委科学技术项目(KJQN202000717)资助

辆传感器实现行人目标轨迹高精度跟踪与预测。其中,相比于视觉、激光雷达等传感器,毫米波雷达通过传输高频电磁波来感知目标,对光照和气候条件具有鲁棒性,能够在全天时、全天候环境下实现高精度的行人感知。此外,得益于调频连续波-多输入多输出(frequency modulated continuous wave-multiple input multiple output, FMCW-MIMO)技术的快速发展,车载毫米波雷达^[4]避免了大规模天线阵列的使用,具备低成本、小尺寸、高精度等优点。因此,本文拟采用 FMCW-MIMO 毫米波雷达实现城市道路交通场景行人目标轨迹跟踪及预测。

尽管 FMCW-MIMO 毫米波雷达具有独特的优势,但在城市复杂道路环境下,可靠的行人目标轨迹跟踪及预测仍然具有挑战性。首先,行人的雷达散射截面积(radar cross section, RCS)约为-5 dBsm,远小于道路基础设施、其他车辆等强反射目标的 RCS,导致行人目标雷达回波信号能量微弱,输出信噪比(signal-to-noise ratio, SNR)低,很容易被强背景杂波淹没^[5]。其次,来自地面、建筑物等多径回波信号叠加在一起会形成多径衰落,造成行人回波信号能量降低^[6],进一步给微弱行人目标跟踪带来困难。同时,行人运动方式灵活多变且行人与社会交互关系复杂导致及时准确地预测行人运动轨迹极具挑战性。

针对低 SNR 场景下微弱行人目标跟踪,传统方法联合阈值检测及数据关联算法,利用杂波及噪声形成虚假点迹的随机无序性,通过引入相关波门,结合卡尔曼滤波(Kalman filter, KF)、粒子滤波(particle filter, PF)等滤波算法实现目标轨迹的提取^[7-8]。然而,由于传统方法引入了阈值判决,在低 SNR 场景下易发生目标漏检现象,导致目标信息丢失,进而恶化微弱行人目标跟踪性能。针对运动目标轨迹预测,按照预测模型复杂度不同,可以将已有方法分为基于浅层学习和深度学习的方法。基于浅层学习的方法将运动学模型与经典机器学习模型相结合,将当前状态传播到未来状态进行轨迹预测,如高斯过程^[9](Gaussian process, GP)、动态贝叶斯网络^[10](dynamic Bayesian network, DBN)等。浅层学习方法为早期行人轨迹预测的发展做出了诸多贡献,但是由于运动特征信息提取不足、模型构建复杂以及大型数据集样本等问题,浅层学习方法难以在较为复杂的场景下进行行人运动建模和高精度轨迹预测。基于深度学习的方法通过复杂数据驱动的深度神经网络结构对变量结果进行潜在学习,从而完成运动目标轨迹预测,如 Social 长短期记忆(long short-term memory, LSTM)网络、StarNet 等^[11-12]。然而,上述方法大多基于递归迭代架构,一条完整的预测轨迹是根据时间顺序逐步处理得到的,无法及时调整学习参数以适应行为的突变。

基于上述分析,本文提出了一种联合递归贝叶斯检测前跟踪和 Transformer 的行人轨迹跟踪-预测一体化方

法,直接利用未经阈值处理的雷达原始频谱数据提取行人目标运动轨迹,在此基础上,采用 Transformer 充分挖掘隐藏在行人运动轨迹中的时空相关性,进而实现微弱行人目标稳定跟踪及可靠预测。

本文提出了一种基于递归贝叶斯检测前跟踪的微弱行人目标轨迹跟踪方法,直接从未经阈值处理的雷达三维原始频谱数据中提取目标运动轨迹,避免了阈值决策过程中出现的信息丢失,提高低 SNR 条件下行人目标轨迹跟踪性能;提出一种基于 Transformer 的端到端行人轨迹预测模型,将轨迹序列作为一个整体进行处理,借助注意力机制对隐藏在长期轨迹中的局部和全局时空特征进行并行编码,避免了在长期特征提取中出现的梯度消失问题;基于 TI 车载 FMCW-MIMO 毫米波雷达,在真实城市道路交通场景下验证了本文所提微弱行人目标轨迹跟踪-预测一体化方法在低 SNR 条件下的有效性。

1 微弱行人目标轨迹跟踪算法设计

本文利用 FMCW-MIMO 毫米波雷达采集行人运动数据,提出的微弱行人目标跟踪-预测一体化方法主要包括基于递归贝叶斯检测前跟踪的行人轨迹跟踪和基于 Transformer 的行人轨迹预测,系统框图如图 1 所示。前者直接利用由三维快速傅里叶变换(three-dimensional fast Fourier transform, 3D FFT)后的非阈值距离-速度-角度观测数据来获得微弱行人目标的当前轨迹。在此基础上,后者借助注意力机制利用行人运动的时空依赖性完成轨迹预测任务。

1.1 FMCW-MIMO 雷达信号处理

在 FMCW 雷达中,发射信号由波形发生器产生,即 $S_t(t_f) = \exp\left[j2\pi\left(f_c t_f + \frac{s_c t_f^2}{2}\right)\right]$ 。对于具有 N_t 个发射天线和 N_r 个接收天线的 MIMO 雷达,可以合成具有 $N_t N_r$ 个天线的虚拟阵列,经过去斜处理过后此虚拟阵列接收行人目标回波信号可以近似表示为:

$$S_{beat}(t_f, t_s, n_a) = A_0 \exp\left[-j2\pi\left(\frac{2R_0}{\lambda} + \frac{2s_c R_0}{c} t_f + \frac{2v}{\lambda} t_s + \frac{d \sin \theta_a}{\lambda} n_a\right)\right] \quad (1)$$

式中: $t_f = n_r T_s \in [0, T_c]$ 表示快时间, $n_r = 0, 1, 2, \dots, N_r - 1$ 表示快时间采样序号, N_r 表示采样总数, T_s 表示采样间隔, T_c 表示脉冲重复间隔, f_c 表示载波频率, $s_c = B/T_c$ 表示调频斜率, B 表示信号带宽, $t_s = n_d T_c$ 表示慢时间, $n_d = 0, 1, 2, \dots, N_d - 1$ 为发射 Chirp 序号, N_d 为发射 Chirp 总数; $n_a = 0, 1, 2, \dots, N_a - 1$ 表示用于角度估计的虚拟阵元天线序号, $N_a = N_t N_r$ 表示虚拟阵元总数; A_0 为信号幅度; R_0 为目标初始距离; $\lambda = c/f_c$ 表示信号波长, c 为光速; v 为目标与雷达之间的相对速度; $d = \lambda/2$ 表示虚拟阵元间隔; θ_a 表示目标角度。

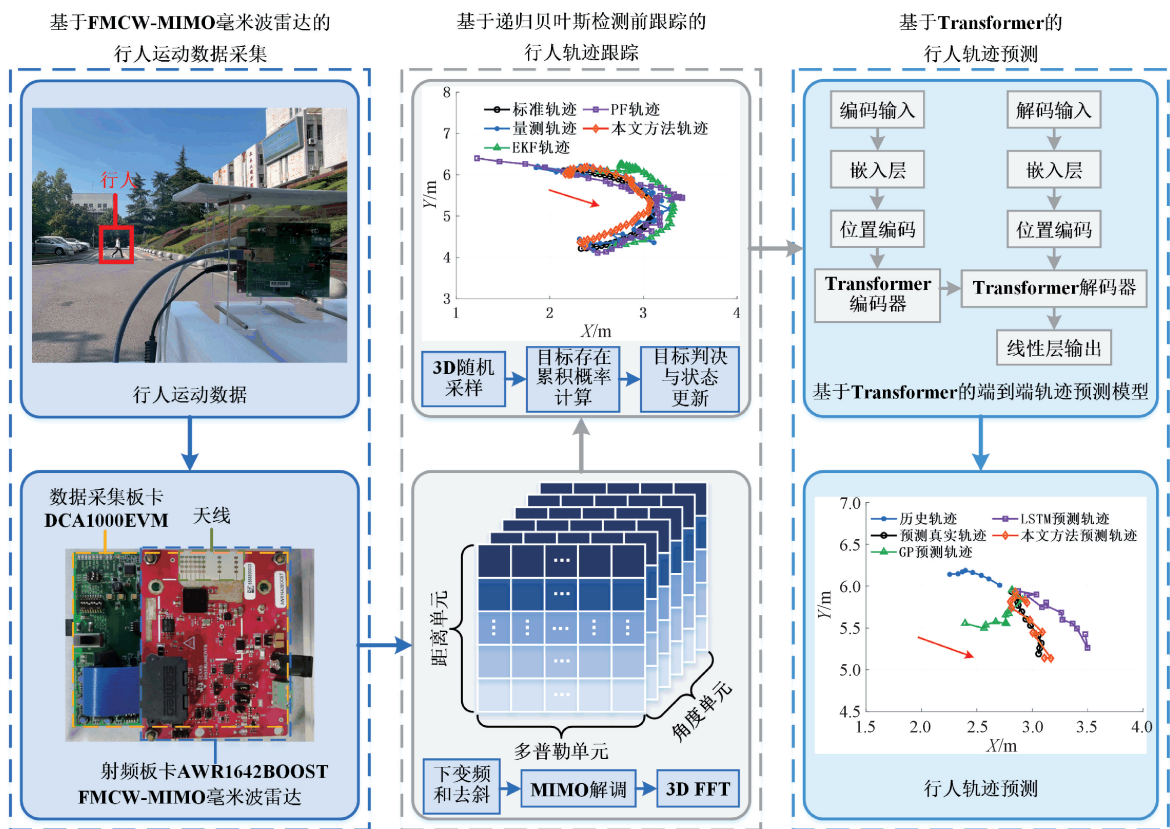


图1 基于毫米波雷达的微弱行人轨迹跟踪-预测一体化系统框图

Fig. 1 System diagram of integrated trajectory tracking and prediction for weak pedestrian with millimeter wave radar

对式(1)依次沿 t_f 、 t_s 和 n_a 维度进行 3D FFT, 得到:

$$\mathcal{F}_n(f_r, f_d, f_a) = A \text{sinc} \left[\pi T_c \left(\frac{2s_c r_n}{c} - f_r \right) \right] \times \text{sinc} \left[\pi N_d T_c \left(\frac{2v_n}{\lambda} - f_d \right) \right] \text{sinc} \left[\pi N_a \left(\frac{d \sin \theta_n}{\lambda} - f_a \right) \right] \quad (2)$$

式中: $\mathcal{F}_n(f_r, f_d, f_a)$ 表示第 n 帧 3D FFT 结果; A 表示复常数; $\text{sinc}(\pi x) = \frac{\sin(\pi x)}{\pi x}$ 表示 sinc 函数; r_n 、 v_n 和 θ_n^a 分别表示第 n 帧距离、速度和角度; f_r 、 f_d 和 f_a 分别表示距离、多普勒和角度频率变量。

从式(2)可以看出, 当目标存在时, $\mathcal{F}_n(f_r, f_d, f_a)$ 在 $f_r = 2s_c r_n / c$, $f_d = 2v_n / \lambda$ 和 $f_a = \sin \theta_n^a / 2$ 处具有峰值, 传统方法通过阈值决策实现目标检测, 很容易导致误检和漏检, 尤其是在低 SNR 环境中的微弱目标检测, 例如微弱行人目标, 会显著降低检测和跟踪性能。因此, 将通过直接使用 3D FFT 后的非阈值雷达原始数据, 并基于递归贝叶斯检测前跟踪的微弱目标轨迹跟踪算法, 以获得低 SNR 条件下行人的运动轨迹。

1.2 基于递归贝叶斯检测前跟踪的行人轨迹提取

行人状态模型表示为:

$$\Phi_{n+1} = \mathcal{S} \Phi_n + \mathcal{N} \mathcal{P}_n \quad (3)$$

式中: $\Phi_n = [x_n, v_x, y_n, v_y]^T$, x_n 和 y_n 分别表示第 n 帧目标

在 X 轴和 Y 轴上的距离坐标, v_x 和 v_y 分别表示第 n 帧目

标沿 X 轴和 Y 轴的速度; $\mathcal{S} = \begin{bmatrix} 1 & I & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & I \\ 0 & 0 & 0 & 1 \end{bmatrix}$ 表示状态转

移矩阵; I 表示两相邻帧之间的间隔; $\mathcal{N} = \begin{bmatrix} I^2/2 & I & 0 & 0 \\ 0 & 0 & I^2/2 & I \end{bmatrix}^T$ 表示噪声转移矩阵; $\mathcal{P}_n$ 表示第 n 帧过程噪声。

雷达量测模型表示为:

$$|\mathcal{F}_n^{n_r, n_d, n_a}| = A_n \beta^{n_r, n_d, n_a}(\Phi_n) + \alpha_n^{n_r, n_d, n_a} \quad (4)$$

式中: $\mathcal{F}_n^{n_r, n_d, n_a}$ 表示在分辨率单元 (n_r, n_d, n_a) 处第 n 帧 3D FFT 结果; $|\cdot|$ 表示取信号幅值; A_n 表示第 n 帧信号幅度; $\beta^{n_r, n_d, n_a}(\Phi_n) = \exp \{ -1/2 [\varepsilon_r (r_n - r_n)^2 / N_r + \varepsilon_d (v_n - v_n)^2 / N_d + \varepsilon_a (\theta_n - \theta_n^a)^2 / N_a] \}$ 表示在分辨率单元 (n_r, n_d, n_a) 处相对于 Φ_n 的指数衰减模型, ε_r 、 ε_d 和 ε_a 分别表示距离、多普勒和角度维度的标准偏差, r_n 、 v_n 和 θ_n^a 分别表示在分辨率单元 (n_r, n_d, n_a) 处的距离、速度和角度变量, $r_n = \sqrt{x_n^2 + y_n^2}$, $v_n = (x_n v_x + y_n v_y) / r_n$, $\theta_n^a = \arctan(x_n / y_n)$; $\alpha_n^{n_r, n_d, n_a}$ 表示在分辨率单元 (n_r, n_d, n_a) 处第 n 帧量测噪声。

由于没有基于阈值的目标检测,因此在当前的原始雷达数据中是否存在目标是不确定的。因此,本文引入目标存在状态 $\Xi_n = \{0, 1\}$ 来指示微弱行人目标是否存在于第 n 帧数据中。如果目标存在,则 $\Xi_n = 1$, 否则 $\Xi_n = 0$ 。此外,基于 Markov 状态转移矩阵 $\Psi = \begin{bmatrix} 1 - P_d & P_b \\ P_d & 1 - P_b \end{bmatrix}$, 两相邻帧之间目标存在概率的递归关系可以表示为:

$$\begin{bmatrix} P(\Xi_{n+1} = 1) \\ P(\Xi_{n+1} = 0) \end{bmatrix} = \begin{bmatrix} 1 - P_d & P_b \\ P_d & 1 - P_b \end{bmatrix} \begin{bmatrix} P(\Xi_n = 1) \\ P(\Xi_n = 0) \end{bmatrix} \quad (5)$$

式中: $P_b = P(\Xi_{n+1} = 1 | \Xi_n = 0)$ 表示目标出生概率; $P_d = P(\Xi_{n+1} = 0 | \Xi_n = 1)$ 表示目标死亡概率。

将目标存在状态与递归贝叶斯理论结合,第 $n + 1$ 帧数据更新过程中目标状态的后验概率可以估计为:

$$P(\Phi_{n+1}, \Xi_{n+1} = 1 | \Omega_{1:n+1}) \propto P(\mathcal{F}_{n+1} | \Phi_{n+1}, \Xi_{n+1} = 1) \times P(\Phi_{n+1}, \Xi_{n+1} = 1 | \Omega_{1:n}) \quad (6)$$

式中: $\Omega_{1:n} = \{\mathcal{F}_1, \mathcal{F}_2, \dots, \mathcal{F}_n\}$ 表示 n 帧雷达量测; $P(\mathcal{F}_{n+1} | \Phi_{n+1}, \Xi_{n+1} = 1)$ 表示第 $n + 1$ 帧似然函数; $P(\Phi_{n+1}, \Xi_{n+1} = 1 | \Omega_{1:n})$ 表示第 $n + 1$ 帧数据预测过程中目标状态的先验概率。

虽然递归贝叶斯方法在理论上可以获得式(6)中目标状态后验概率的最优估计,但是由于系统的非线性和积分运算的复杂性,无法在实践中应用。因此,本文引入了序贯蒙特卡罗(sequential Monte Carlo, SMC)方法^[13]来近似后验概率密度函数。

首先,在第 n 帧数据中随机产生 N_s 个粒子用于初始化。其次,借助式(5)预测第 $n + 1$ 帧粒子状态。然后,第 $n + 1$ 帧预测粒子被分为新生粒子、继续粒子和死亡粒子。随后,结合第 $n + 1$ 帧雷达量测信息,更新每个粒子的重要性权重,得到:

$$\omega_{n+1}^p = \begin{cases} \mathcal{L}^p(\mathcal{F}_{n+1} | \Phi_{n+1}, \Xi_{n+1} = 1), & \Xi_{n+1}^p = 1 \\ 1, & \Xi_{n+1}^p = 0 \end{cases} \quad (7)$$

式中: ω_{n+1}^p 表示第 $n + 1$ 帧数据第 p 个粒子的更新权重; $\mathcal{L}^p(\mathcal{F}_{n+1} | \Phi_{n+1}, \Xi_{n+1} = 1)$ 表示目标存在时第 $n + 1$ 帧数据第 p 个粒子的似然比。

$$\mathcal{L}^p(\mathcal{F}_{n+1} | \Phi_{n+1}, \Xi_{n+1} = 1) = \exp\{A_{n+1}^p \beta^{n_r, n_d, n_a}(\Phi_{n+1}^p) / 2\delta^2 \times [2 \mathcal{F}_{n+1}^{n_r, n_d, n_a} - A_{n+1}^p \beta^{n_r, n_d, n_a}(\Phi_{n+1}^p)]\} \quad (8)$$

式中: δ^2 表示高斯分布方差。通过归一化和重采样第 $n + 1$ 帧粒子权重,计算出目标存在累积概率为 $P_{\Xi_{n+1}} = \sum_{j=1}^{N_s} \Xi_{n+1}^j / N_s$ 。最后,如果 $P_{\Xi_{n+1}}$ 大于预设阈值,则判断第 $n + 1$ 帧目标存在,并计算粒子状态期望值以提取目标轨迹信息,即 $\tilde{\Phi}_{n+1} = (\sum_{j=1}^{N_s} \Xi_{n+1}^j \Phi_{n+1}^j) / N_{\Xi_{n+1}}$, 其中 $\tilde{\Phi}_{n+1}$ 是第

$n + 1$ 帧估计的目标状态, $N_{\Xi_{n+1}}$ 表示目标存在状态为 1 的粒子数。

2 微弱行人目标轨迹预测算法设计

2.1 基于 Transformer 的行人轨迹预测模型

为了在低 SNR 环境下获得高精度轨迹预测性能,采用第 1 节获得的微弱行人目标跟踪轨迹作为基于 Transformer 的端到端模型的输入序列来完成行人轨迹预测任务。在该模型中,编码器分析行人运动轨迹的时空相关性,并将轨迹特征转化为隐状态;解码器通过从输入序列中提取识别结果来生成行人的未来轨迹。

行人当前轨迹观测序列 $O = \{\Phi'_{-N}, \Phi'_{-N+1}, \dots, \Phi'_{-1}\}$, $O \in \mathbb{R}^{N \times 2}$, 其中, $\Phi'_n = [x_n, y_n]^T$ 表示 n 时刻观测位置, N 表示观测序列长度。为了便于模型处理输入序列,通过具有权重矩阵 $W_0 \in \mathbb{R}^{2 \times D_E}$ 的线性投影层将其嵌入到更高的 D_E 维空间,即 $O' = OW_0$, $O' \in \mathbb{R}^{N \times D_E}$ 。此外,为了从整个输入序列中识别节点的顺序,引入正弦和余弦函数值作为标记来识别不同的节点,位置编码后输入序列表示为 $O_0 = O' + P_C$, 其中, $P_C = \{p_{n,d}, n = 1, 2, \dots, N, d = 1, 2, \dots, D_E\}$ 表示位置编码向量, $p_{n,d}$ 表示为:

$$p_{n,d} = \begin{cases} \sin\left(\frac{n}{10\,000^{d/D_E}}\right), & d = 2, 4, 6, \dots \\ \cos\left(\frac{n}{10\,000^{d/D_E}}\right), & d = 1, 3, 5, \dots \end{cases} \quad (9)$$

1) 编码器

编码序列 O_0 被输送到具有 M 个相同层的堆栈组成的编码器中。每层由一个多头自注意力子层和一个位置前馈子层组成。此外,为了解决破碎梯度问题并提高深度模型的收敛能力,对这两个子层中的每一个子层使用残差连接,然后进行层归一化。编码器每层的输出序列可以表示为:

$$O_1 = \text{MultiHead}(O_E^m, O_E^m, O_E^m) \quad (10)$$

$$O_2 = \text{LayerNorm}(O_1 + O_E^m) \quad (11)$$

$$O_3 = \text{ReLU}(O_2 W_1 + b_1) W_2 + b_2 \quad (12)$$

$$O_E^{m+1} = \text{LayerNorm}(O_2 + O_3) \quad (13)$$

式中: $O_E^m, m = 0, 1, \dots, M - 1$ 表示第 m 个编码器层的输出序列(其中, $O_E^0 = O_0$); W_1, W_2 和 b_1, b_2 分别表示位置前馈子层训练权重和偏置。

根据多头自注意力机制^[14],式(10)表示多头自注意力子层,并且 3 个输入相同用于从不同表示子空间中获来自注意力。式(11)和(13)表示归一化子层,在特征维度上归一化先前子层的输出以提高模型泛化能力。式(12)表示位置前馈网络,包括整流线性单元(rectified

linear unit, ReLU) 子层和两个全连接子层, 利用 ReLU 激活的两个线性变换来改善模型的非线性能力。

2) 解码器

解码器的结构类似于编码器, 也由 M 个相同层的堆栈组成。除了包含每个编码器层中的两个子层外, 解码器还引入第 3 个子层, 该子层对编码器堆栈的输出执行多头注意力, 处理当前行人轨迹的信息。与编码器类似, 对每一个子层使用残差连接, 然后进行层归一化。此外, 对解码器堆栈中的多头自注意力子层进行序列掩码操作, 以防止关注未来轨迹位置信息。这种掩码操作, 再加上输出嵌入偏移一个位置, 确保了对位置 n 处的预测只能取决于小于 n 的位置处的已知输出。

解码是自回归过程, 并通过 N' 步迭代生成行人的轨迹预测结果 $\mathbf{Y}_p^{N'} = \{\mathbf{y}_p^1, \mathbf{y}_p^2, \dots, \mathbf{y}_p^{N'}\}$, 其中 $\mathbf{y}_p^{n'}$ 是未来 n' 时刻的轨迹向量。在迭代过程中, 在 n' 时刻, 状态 Φ'_0 和先前预测的轨迹序列 $\mathbf{Y}_p^{n'-1} = \{\mathbf{y}_p^1, \mathbf{y}_p^2, \dots, \mathbf{y}_p^{n'-1}\}$ 首先级联为 $\Pi = \{\Phi'_0, \mathbf{y}_p^1, \mathbf{y}_p^2, \dots, \mathbf{y}_p^{n'-1}\}$ 。然后, 与编码器类似, 采用嵌入和位置编码来获得具有节点标记的高维输入序列。解码器每层的输出序列可以表示为:

$$\mathbf{O}_5 = \text{MultiHead}(\text{Masked}(\mathbf{O}_D^m, \mathbf{O}_D^m, \mathbf{O}_D^m)) \quad (14)$$

$$\mathbf{O}_6 = \text{LayerNorm}(\mathbf{O}_5 + \mathbf{O}_D^m) \quad (15)$$

$$\mathbf{O}_7 = \text{MultiHead}(\mathbf{O}_6, \mathbf{O}_E^M, \mathbf{O}_E^M) \quad (16)$$

$$\mathbf{O}_8 = \text{LayerNorm}(\mathbf{O}_7 + \mathbf{O}_6) \quad (17)$$

$$\mathbf{O}_9 = \text{ReLU}(\mathbf{O}_8 \mathbf{W}_3 + \mathbf{b}_3) \mathbf{W}_4 + \mathbf{b}_4 \quad (18)$$

$$\mathbf{O}_D^{m+1} = \text{LayerNorm}(\mathbf{O}_9 + \mathbf{O}_8) \quad (19)$$

式中: $\mathbf{O}_D^m$ 表示第 m 个解码器层的输出序列 (其中, $\mathbf{O}_D^0 = \Pi \mathbf{W}_0 + \mathbf{P}_C$); $\mathbf{W}_3$ 、 $\mathbf{W}_4$ 和 $\mathbf{b}_3$ 、 $\mathbf{b}_4$ 分别表示位置前馈子层训练权重和偏置。根据式 (14) ~ (19) 可知, n' 时刻的预测轨迹序列估计为:

$$\mathbf{Y}_p^{N'} = \mathbf{O}_D^M \mathbf{W}_5 + \mathbf{b}_5 = \{\mathbf{y}_p^1, \mathbf{y}_p^2, \dots, \mathbf{y}_p^{N'}\} \quad (20)$$

式中: $\mathbf{W}_5$ 和 $\mathbf{b}_5$ 分别表示线性层训练权重和偏置。

2.2 实施细节

本文在 ETH^[15] 和 UCY^[16] 两个公开行人轨迹数据集上训练和测试模型, 在雷达采集行人轨迹数据集上推断预测结果。将公开数据集与本文所采集的雷达数据集相结合, 目的是为了增加本文预测模型训练数据的多样性, 从而提高微弱行人目标轨迹预测的准确性和可靠性, 同时运用不同的数据集也能体现本文预测模型在不同数据集上的泛化能力。ETH 数据集有 ETH 和 HOTEL 两个场景, UCY 数据集有 UNIV、ZARA1 和 ZARA2 这 3 个场景, 雷达数据集有一个交通场景。训练和测试过程采用留一法 (leave-one-out, LOO), 使用 ETH 和 UCY 上 4 个数据集进行训练, 在剩余一个数据集上进行测试。此外, 每 0.4 s 从数据中采样一次, 以获得轨迹。观测行人

8 frame, 并获得接下来 12 frame 的真实轨迹数据, 以评估预测。本文采用平均位移误差 (average displacement error, ADE) 和最终位移误差 (final displacement error, FDE)^[17] 两种指标来评估模型性能和预测结果。

本文基于 Transformer 的行人轨迹预测模型采用的参数包括隐藏特征维度 $D_E = 512$ 、编码器和解码器层数 $M = 6$ 、自注意力头数为 8。采用 L2 损失, 并使用 Adam 优化器通过反向传播训练网络, 前 5 轮为线性预热阶段, 之后学习率递减, dropout 值为 0.1, 训练轮数为 240。ADE、FDE 随训练轮数的变化情况如图 2 所示, ADE、FDE 随训练轮数逐渐下降直至收敛, 最终 ADE、FDE 分别为 0.445、0.965 m。

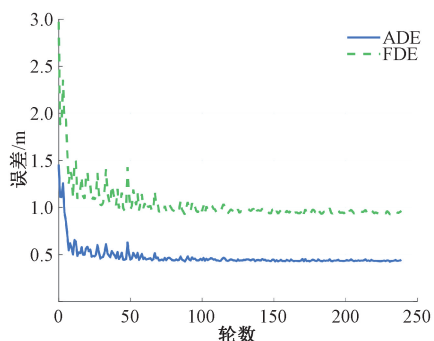


图 2 ADE、FDE 随训练轮数的变化情况

Fig. 2 ADE、FDE with training epochs varying

3 实验与结果分析

3.1 实验设备及场景

基于 TI 公司的 AWR1642 FMCW-MIMO 雷达, 在真实道路场景中验证行人轨迹跟踪-预测一体化方法的有效性。AWR1642 FMCW-MIMO 雷达与 DCA1000EVM 数据采集板卡相结合, 能够获得包含雷达回波所有信息的原始 ADC 数据。AWR1642 毫米波雷达中心频点为 77 GHz, 具有 4 GHz 可用带宽、2 个发射天线和 4 个接收天线。真实城市道路交通场景如图 3 所示, 实验中雷达处于静止状态, 行人相对于雷达运动。需要说明的是, 由于雷达量测为相对信息, 因此只要雷达与行人目标发生相对运动即可实现行人目标跟踪, 从而当雷达处于车载平台上运动时也可以利用本文方法得到行人目标稳定的跟踪结果。

3.2 微弱行人目标轨迹跟踪-预测实验结果

1) 不同距离下轨迹跟踪-预测实验结果

为了验证本文轨迹跟踪-预测一体化方法在低 SNR 环境下的有效性, 在行人目标距离雷达 6 和 30 m 处进行了两组实验, 实验结果如图 4 所示。行人目标与雷达距

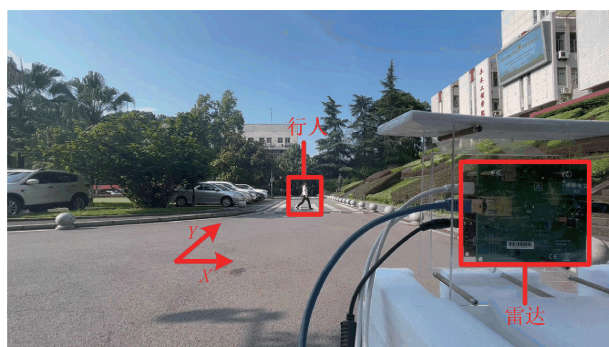


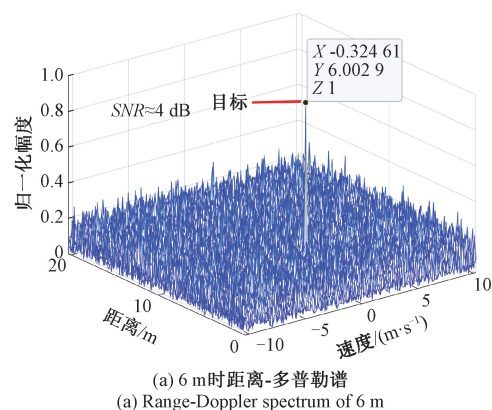
图 3 实验场景

Fig. 3 Experimental scene

离 6 m 时距离-多普勒谱如图 4(a) 所示, 此时行人目标回波信号能量远大于干扰噪声能量, 目标峰值明显,

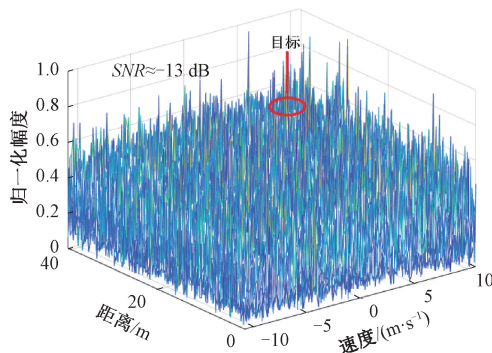
$SNR = 10 \log \left(\frac{P_s}{P_n} \right) - 10 \lg N_d \approx 4 \text{ dB}$, 其中 $P_N =$

$\frac{1}{N_r N_d} \left[\sum_{n_r} \sum_{n_d} P(n_r, n_d) - P_s \right]$ 表示噪声功率, 即除目标所在峰值外所有距离-多普勒单元功率的平均值, P_s 表示信号功率, $10 \lg N_d$ 表示多普勒处理增益; 而当目标与雷达距离增加到 30 m 时, 目标回波信号距离-多普勒谱如



(a) 6 m 时距离-多普勒谱

(a) Range-Doppler spectrum of 6 m



(b) 30 m 时距离-多普勒谱

(b) Range-Doppler spectrum of 30 m

图 4 不同距离行人回波距离-多普勒谱

Fig. 4 Range-Doppler spectrum with different distances

图 4(b) 所示, 由于行人目标回波信号能量衰减与距离 4 次方成正比, 导致行人目标回波信号能量远低于干扰噪声能量, 此时 $SNR \approx -13 \text{ dB}$, 行人目标完全被噪声所淹没。

不同距离的轨迹跟踪结果如图 5(a)、(d) 所示。为了验证跟踪性能, 本文引入雷达量测轨迹与通过两种传统跟踪方法, 即扩展卡尔曼滤波 (extended KF, EKF) 和 PF, 获得的跟踪轨迹作为对比。可以看出, 行人目标与雷达距离为 6 m 时, 雷达量测、EKF、PF 和本文方法具有类似的轨迹跟踪性能。然而随着距离增加到 30 m, 雷达量测、EKF 和 PF 的跟踪性能严重恶化, 因为它们在很大程度上依赖于传统目标检测算法中的阈值决策过程, 在目标与雷达距离为 30 m 条件下会引入漏检并产生虚警, 严重恶化行人目标跟踪结果。相比于雷达量测、EKF 和 PF, 本文方法在 6 和 30 m 都能取得较好的微弱行人目标跟踪结果。不同距离下跟踪误差累积分布函数 (cumulative distribution functions, CDFs) 如图 5(b)、(e) 所示。可以看出, 当目标与雷达距离为 6 m 时, 雷达量测、EKF、PF 和本文方法具有相近的跟踪结果。然而, 当目标与雷达距离增加到 30 m 时, 本文方法跟踪结果距离估计误差小于 2 m 的概率大于 0.9, 远胜于雷达量测、EKF 和 PF。不同距离的轨迹预测结果如图 5(c)、(f) 所示, 本文将 8 frame 基于递归贝叶斯的跟踪轨迹作为观测数据预测未来 12 frame 的轨迹。同时, 为了验证本文轨迹预测模型的有效性, 本文选择了传统机器学习预测算法 GP 和经典深度学习预测算法 LSTM 作为对比。可以看出, 相比于 GP 和 LSTM, 本文方法预测的轨迹更接近真实轨迹。

表 1 和 2 为不同距离下的预测 ADE 和 FDE。从表中可知, 本文方法的预测 ADE 和 FDE 比 GP 和 LSTM 的都要小, 同样说明本文轨迹预测方法的性能更好。相比之下, GP 的轨迹预测性能最差, 这是由于 GP 是一个潜在变量模型, 包括具有相关动力学的低维潜在空间, 以及从低维潜在空间到高维观测空间的映射, 进而从高维运动捕捉数据中学习人体运动模型。但是由于目标的未来轨迹是不确定的, 导致低维潜在空间到高维观测空间的映射关系建立困难。LSTM 的轨迹预测效果不如本文基于 Transformer 的轨迹预测方法, 因为 LSTM 在开始自回归预测之前顺序处理观测结果, 而本文基于 Transformer 的轨迹预测方法查看所有可用的观测结果, 并根据注意力机制对其进行加权。综上所述, 与传统方法相比, 本文方法更适合于低 SNR 环境下微弱行人目标轨迹跟踪和预测。

2) 复杂轨迹场景下跟踪-预测实验结果

为了验证本文轨迹跟踪-预测一体化方法在复杂轨迹场景下的有效性, 行人相对于雷达做曲线运动, 实验结果如图 6 所示。图 6(a) 为曲线运动的轨迹跟踪结果,

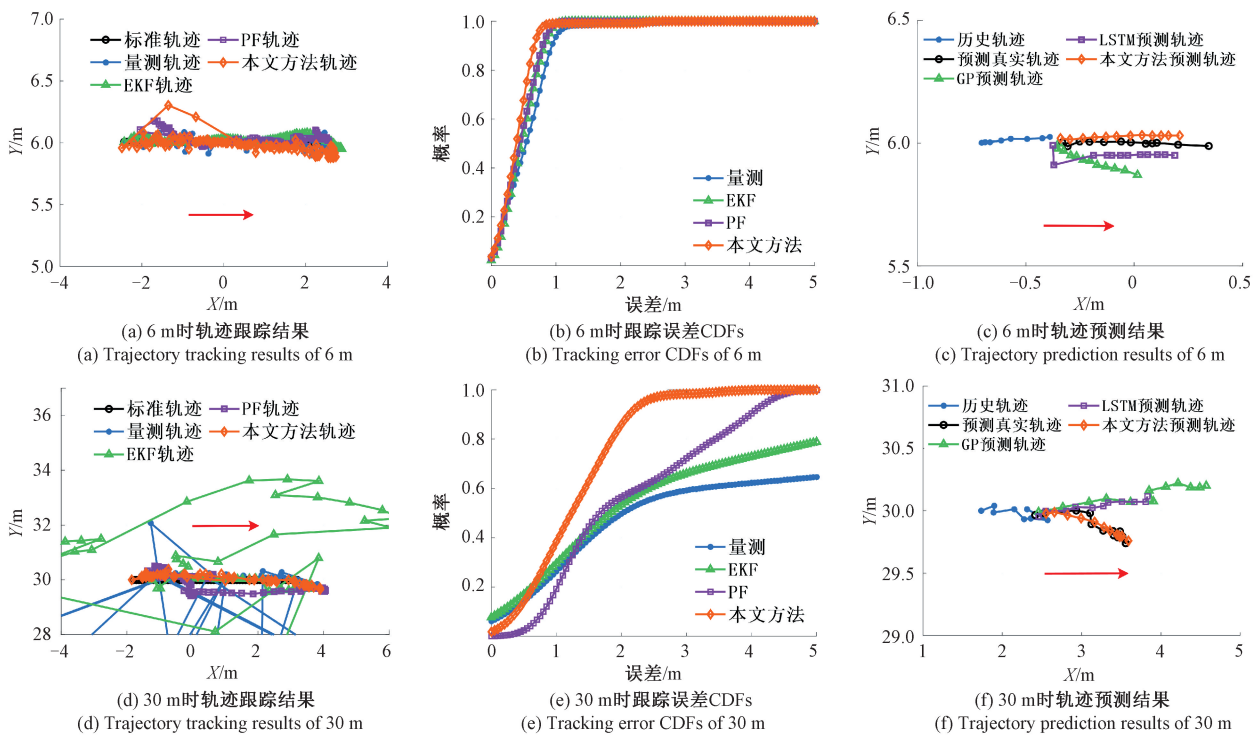


图5 不同距离轨迹跟踪和预测结果

Fig. 5 Trajectory tracking and prediction results with different distances

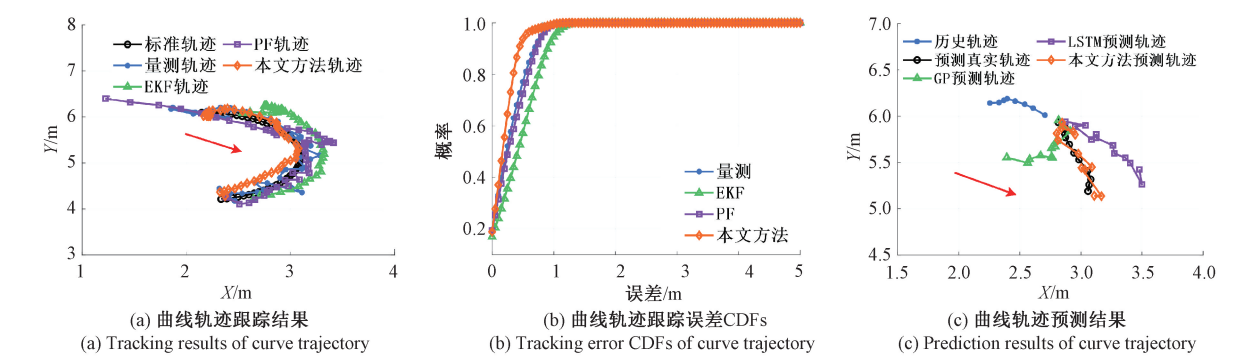


图6 复杂轨迹场景下跟踪和预测结果

Fig. 6 Tracking and prediction results in complex trajectory scenarios

表1 6 m 时预测 ADE 和 FDE

Table 1 Prediction ADE and FDE of 6 m			
模型	GP	LSTM	本文方法
ADE	0.175	0.076	0.043
FDE	0.349	0.162	0.141

表2 30 m 时预测 ADE 和 FDE

Table 2 Prediction ADE and FDE of 30 m			
模型	GP	LSTM	本文方法
ADE	0.550	0.325	0.084
FDE	1.150	0.528	0.164

图 6(b) 为对应的跟踪误差 CDFs, 实验结果同上, 本文基于递归贝叶斯检测前跟踪的轨迹跟踪方法通过联合处理检测和跟踪过程, 其跟踪性能优于 EKF、PF 这两种传统跟踪方法, 并且在复杂轨迹场景下仍然保持一致的优势。图 6(c) 为曲线运动的轨迹预测结果, 表 3 为曲线运动的预测 ADE 和 FDE, 实验结果同上, 本文基于 Transformer 的行人轨迹预测模型由于注意力机制的自回归特性和长期预测优势, 使得模型能够更有效地利用已有轨迹, 进一步挖掘数据间的内部信息, 并通过迭代机制不断修正预测轨迹, 在曲线的复杂轨迹场景下仍然可以获得准确的未来轨迹, 其预测性能好于 GP 和 LSTM 这两种经典预测方法。综上所述, 本文方法仍然

适合于复杂轨迹场景下微弱行人目标轨迹跟踪和预测。

表 3 曲线轨迹预测 ADE 和 FDE

Table 3 Prediction ADE and FDE of curve trajectory			
m			
模型	GP	LSTM	本文方法
ADE	0.422	0.273	0.087
FDE	0.759	0.482	0.071

4 讨 论

4.1 SNR 对跟踪-预测结果的影响

由于 SNR 影响轨迹跟踪误差,而跟踪误差影响预测性能,因此在这一部分中,评估了 SNR 对本文方法的轨迹跟踪和预测性能,如图 7 所示。图 7(a)为输入 SNR 在 -30~20 dB 变化时,行人轨迹跟踪的均方根误差 (root-mean-square errors, RMSEs)。可以看出,当 SNR 足够高时,本文方法具有与 EKF 和 PF 接近的跟踪 RMSEs。但

当 $SNR < -10$ dB 时,雷达量测的跟踪 RMSEs 随着 SNR 的降低而迅速增加。因此,高度依赖雷达量测的 EKF 和 PF 的跟踪性能也迅速恶化。然而,本文方法的跟踪 RMSEs 直到输入 $SNR < -20$ dB 时才显著增加。因此,通过直接使用 3D FFT 后的非阈值雷达原始数据,本文方法在低 SNR 环境下具有优于 EKF 和 PF 的跟踪能力。图 7(b)为受跟踪 RMSEs 影响的预测 ADE 和 FDE 的结果。可以看出,随着跟踪 RMSEs 的增加,每个模型的预测误差逐渐增大。但与 GP 和 LSTM 相比,本文方法在每个跟踪 RMSEs 下都具有更小的预测 ADE 和 FDE。

4.2 观测长度对预测结果的影响

由于观测长度影响轨迹预测结果,在这一部分中,研究了当输入观测长度发生变化时模型评估的性能如表 4 所示。从表 4 可以看出,减少历史轨迹的观测数量会导致模型预测性能显著下降,即使对于像删除单个时间步长的小变化也是如此。然而,与 GP 和 LSTM 相比,本文方法在每个观测长度下都具有一致的优势。

表 4 在不同观测长度下的 ADE 和 FDE

Table 4 ADE and FDE with different observation lengths			
观测长度	GP	LSTM	本文方法
/frame	ADE/m, FDE/m	ADE/m, FDE/m	ADE/m, FDE/m
2	3.535, 4.839	2.138, 2.819	2.137, 2.663
3	3.518, 4.923	2.219, 2.877	2.049, 2.578
4	3.206, 4.412	2.005, 2.645	1.935, 2.456
5	2.617, 3.724	1.414, 2.016	1.336, 1.861
6	2.183, 3.086	1.227, 1.845	1.112, 1.635
7	2.326, 3.428	1.181, 1.762	1.135, 1.651
8	2.101, 3.102	0.916, 1.515	0.836, 1.355

4.3 预测长度对预测结果的影响

由于预测长度影响轨迹预测结果,在这一部分中,比较了 GP、LSTM 和本文方法在预测较长时间上的稳定性,预测长度为 12~32 frame,在不同预测长度下的 ADE、FDE 如表 5 所示。从表 5 可以看出,随着预测长度的增加,每个模型的性能普遍下降。但与 GP 和 LSTM 相比,本文方法在每个预测长度下都具有一致的优势。

4.4 不同轨迹预测方法计算复杂度分析

考虑到车载计算机算力有限,本文首先在系统设计上,利用二维轨迹跟踪数据作为预测网络输入进行行人轨迹预测,而不是与传统方法一样将高维距离-多普勒谱或时频谱数据作为网络输入,进而在输入数据层面降低了本文轨迹预测方法的计算复杂度。本文方法、LSTM 和 GP 的计算复杂度分别为 $O(l^2 \cdot d)$ 、 $O(l \cdot d^2)$ 和 $O(l^3)$ 。可以看到,当序列长度 l 大于隐藏层特征维度 d 时,本文

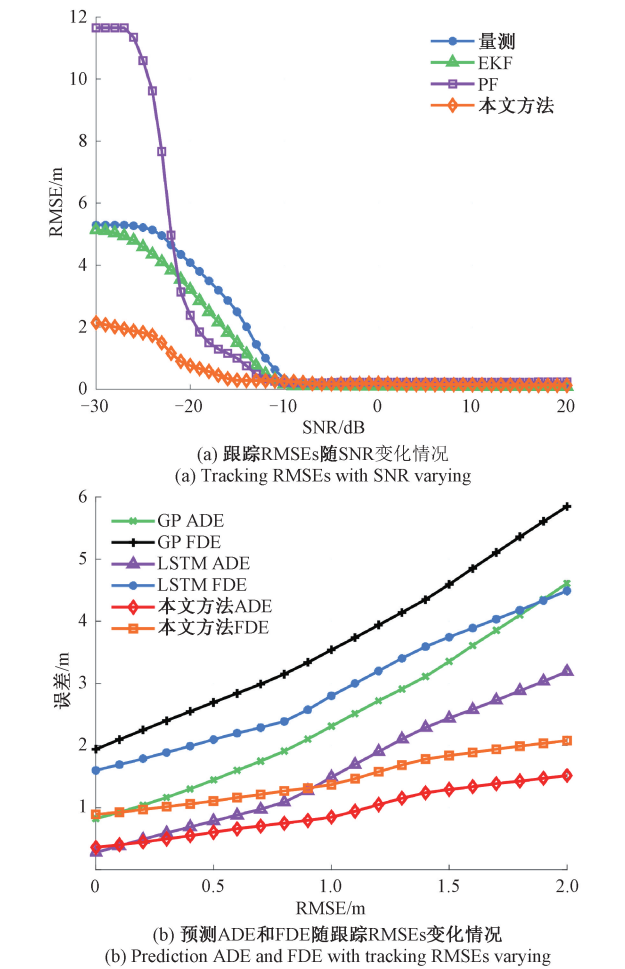


图 7 SNR 对轨迹跟踪和预测的影响
Fig. 7 The impact of SNR on trajectory tracking and prediction

表 5 在不同预测长度下的 ADE 和 FDE
Table 5 ADE and FDE with different prediction lengths

观测长度 /frame	GP	LSTM	本文方法
	ADE、FDE/m	ADE、FDE/m	ADE、FDE/m
12	2. 101、3. 102	0. 916、1. 515	0. 836、1. 355
16	3. 076、4. 045	1. 285、2. 546	1. 075、1. 946
20	3. 494、5. 697	1. 777、3. 794	1. 397、2. 694
24	3. 883、7. 651	2. 435、5. 374	1. 785、3. 554
28	4. 498、9. 979	3. 217、7. 275	2. 397、4. 885
32	5. 205、9. 422	4. 286、9. 786	3. 106、4. 316

所提方法的计算复杂度比 GP 小、比 LSTM 大;当序列长度 l 小于隐藏层特征维度 d 时,本文所提算法计算复杂度比 LSTM 小、比 GP 大。相比于 LSTM,虽然本文方法模型内序列长度的平方项增加了本文方法对于长时间序列预测的计算复杂度,但本文方法可以更好地捕获轨迹跟踪序列长距离依赖关系,从而得到更精确的轨迹预测结果。同样,GP 在长序列学习任务中计算复杂度高,由于 GP 属于无参数模型,不适用于解决非稳态数据,因此本文方法的计算复杂度和轨迹预测精度均优于 GP。同时,本文方法适用于并行计算,可以利用图形处理器进行运算,进一步降低本文方法的计算复杂度,增加算法的实用性。

5 结 论

本文提出了一种低信噪比场景 FMCW-MIMO 雷达微弱行人轨迹跟踪及预测一体化方法,主要特点如下:1) 为了提高低信噪比场景微弱行人目标跟踪性能,本文基于目标回波信号多帧之间的时空相关特性,提出了基于递归贝叶斯检测前跟踪的微弱行人目标跟踪方法,既实现了微弱行人目标连续多帧回波信号能量的累积,又避免了传统阈值检测带来的信息丢失;2) 为了应对行人运动的随机突变性,借助注意力机制,提出了一种基于 Transformer 的端到端轨迹预测模型,同时实现对隐藏在长期轨迹中的局部和全局特征进行编码;3) 本文方法通过 FMCW-MIMO 车载毫米波雷达进行了验证,并分析了 SNR、观测轨迹时长、预测轨迹时长等因素对行人目标轨迹预测的影响。实验结果表明,本文方法可以在低 SNR 条件下获得优越的行人目标轨迹跟踪及预测性能。同时,考虑到在实际应用中,自动驾驶车辆需要通过融合视觉、雷达等多传感器感知数据,提升行人轨迹跟踪与预测性能。因此,未来将进一步结合视觉感知传感器,设计基于视觉和雷达多源数据融合的微弱行人目标感知系统,实现自动驾驶车辆全天时、全天候环境下高精度行人感知。

参考文献

[1] CHEN H, LIU Y, ZHAO B, et al. Vision-based real-time online vulnerable traffic participants trajectory prediction for autonomous vehicle[J]. IEEE Transactions on Intelligent Vehicles, 2023, 8(3): 2110-2122.

[2] BU F, LE T, DU X, et al. Pedestrian planar LiDAR pose (PPLP) network for oriented pedestrian detection based on planar LiDAR and monocular images[J]. IEEE Robotics and Automation Letters, 2020, 5 (2): 1626-1633.

[3] ENGELS F, HEIDENREICH P, WINTERMANTEL M, et al. Automotive radar signal processing: Research directions and practical challenges[J]. IEEE Journal of Selected Topics in Signal Processing, 2021, 15 (4): 865-878.

[4] TAN B, MA Z, ZHU X, et al. 3-D object detection for multiframe 4-D automotive millimeter-wave radar point cloud[J]. IEEE Sensors Journal, 2023, 23 (11): 11125-11138.

[5] BELGIOVANE D, CHEN C C, CHEN M, et al. 77 GHz radar scattering properties of pedestrians[C]. 2014 IEEE Radar Conference, IEEE, 2014: 0735-0738.

[6] AYDOGDU C, KESKIN M F, CARVAJAL G K, et al. Radar interference mitigation for automated driving: Exploring proactive strategies [J]. IEEE Signal Processing Magazine, 2020, 37(4): 72-84.

[7] 张磊, 彭力. 结合 SURF 与 Kalman 滤波的 CAMShift 跟踪算法[J]. 电子测量与仪器学报, 2017, 31(3): 389-394.

ZHANG L, PENG L. CAMShift tracking algorithm of combined with SURF and Kalman filter[J]. Journal of Electronic Measurement and Instrumentation, 2017, 31(3): 389-394.

[8] 杨傲雷, 金宏宙, 陈灵, 等. 融合深度学习与粒子滤波的移动机器人重定位方法[J]. 仪器仪表学报, 2021, 42(7): 226-233.

YANG AO L, JIN H ZH, CHEN L, et al. Mobile robot relocation method fusing deep learning and particle filtering[J]. Chinese Journal of Scientific Instrument, 2021, 42(7): 226-233.

[9] ELLIS D, SOMMERLADE E, REID I. Modelling pedestrian trajectory patterns with Gaussian processes [C]. 2009 IEEE 12th International Conference on Computer Vision Workshops, IEEE, 2009: 1229-1234.

[10] KOOIJ J F P, SCHNEIDER N, FLOHR F, et al. Context-based pedestrian path prediction [C]. Computer Vision-ECCV 2014: 13th European Conference, Springer, 2014: 618-633.

- [11] WANG R, SONG X, HU Z, et al. Spatio-temporal interaction aware and trajectory distribution aware graph convolution network for pedestrian multimodal trajectory prediction [J]. IEEE Transactions on Instrumentation and Measurement, 2023, 72: 1-11.
- [12] GUAN H, GUO P. Research on pedestrian trajectory prediction by GAN model based on LSTM [C]. 2023 IEEE 3rd International Conference on Power, Electronics and Computer Applications (ICPECA), IEEE, 2023: 1400-1405.
- [13] 方鑫, 朱婧, 黄大荣, 等. 低 SNR 场景下微型无人机跟踪-检测融合方法 [J]. 仪器仪表学报, 2022, 43(4): 79-88.
FANG X, ZHU J, HUANG D R, et al. Integrated tracking and detection of micro UAV under low SNR environment [J]. Chinese Journal of Scientific Instrument, 2022, 43(4): 79-88.
- [14] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [C]. Proceedings of the 31st International Conference on Neural Information Processing Systems, 2017.
- [15] PELLEGRINI S, ESS A, SCHINDLER K, et al. You'll never walk alone: Modeling social behavior for multi-target tracking [C]. 2009 IEEE 12th International Conference on Computer Vision, IEEE, 2009: 261-268.
- [16] LEALTAIXE L, FENZI M, KUZNETSOVA A, et al. Learning an image-based motion context for multiple people tracking [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014: 3542-3549.
- [17] 王海峰, 桑海峰, 王金玉, 等. 结合速度控制的时空图网络行人轨迹预测模型 [J]. 电子测量与仪器学报, 2022, 36(5): 146-154.
WANG H F, SANG H F, WANG J Y, et al. Spatial-temporal graph network with speed control pedestrian trajectory prediction model [J]. Journal of Electronic Measurement and Instrumentation, 2022, 36(5): 146-154.

作者简介



方鑫, 分别在 2012 年和 2015 年于重庆邮电大学获学士学位和硕士学位, 2019 年于电子科技大学获博士学位, 现为西南石油大学讲师, 主要研究方向为多智能体协作感知。

E-mail: fangxin@tom.com

Fang Xin received his B.Sc. degree from Chongqing University of Posts and Telecommunications in 2012, M.Sc. degree from Chongqing University of Posts and Telecommunications in 2015, and Ph.D. degree from University of Electronic Science and Technology of China in 2019. He is currently a lecturer at Southwest Petroleum University. His main research interest includes multi-agent collaborative perception.



何敏, 2020 年于成都理工大学获得学士学位, 现为西南石油大学硕士研究生, 主要研究方向为雷达信号处理。

E-mail: hm2225279736@163.com

He Min received her B.Sc. degree from Chengdu University of Technology in 2020. Now she is a M.Sc. candidate at Southwest Petroleum University. Her main research interest includes radar signal processing.



黄大荣 (通信作者), 2000 年于湖北学院获得学士学位, 2003 年于辽宁大学获得硕士学位, 2006 年于重庆大学获得博士学位, 现为安徽大学教授, 主要研究方向为多智能体协作感知与容错控制。

E-mail: drhuang@cqjtu.edu.cn

Huang Darong (Corresponding author) received his B.Sc. degree from Hubei National Institute in 2000, M.Sc. degree from Liaoning University in 2003 and Ph.D. degree from Chongqing University in 2006. He is currently a professor at Anhui University. His main research interests include multi-agent collaborative perception and fault-tolerant control.



张振源, 分别在 2012 年于黄淮学院获得学士学位, 2015 年和 2019 年于重庆邮电大学获得硕士学位和博士学位, 现为重庆交通大学副教授, 主要研究方向为深度学习及雷达信号处理。

E-mail: zhenyuan@cqjtu.edu.cn

Zhang Zhenyuan received his B.Sc. degree from HuangHuai University in 2012, M.Sc. degree from Chongqing University of Posts and Telecommunications in 2015 and Ph.D. degree from Chongqing University of Posts and Telecommunications in 2019. He is currently an associate professor at Chongqing Jiaotong University. His main research interests include deep learning and radar signal processing.