



行人惯性定位新动态: 基于神经网络的方法、性能与展望

李岩 施忠臣 侯燕青 戚煜华 谢良 陈伟 陈洪波 闫野 印二威

Emerging Trends in Pedestrian Inertial Positioning: Neural Network-based Methods, Performance and Prospects

LI Yan, SHI Zhong-Chen, HOU Yan-Qing, QI Yu-Hua, XIE Liang, CHEN Wei, CHEN Hong-Bo, YAN Ye, YIN Er-Wei

在线阅读 View online: <https://doi.org/10.16383/j.aas.c240221>

您可能感兴趣的其他文章

基于多层BP神经网络的无参考视频质量客观评价

No Reference Video Quality Objective Assessment Based on Multilayer BP Neural Network

自动化学报. 2022, 48(2): 594–607 <https://doi.org/10.16383/j.aas.c190539>

惯性组合导航系统性能评估方法研究进展

Perspectives on Performance Evaluation Method for Inertial Integrated Navigation System

自动化学报. 2022, 48(10): 2361–2373 <https://doi.org/10.16383/j.aas.c210377>

一种基于动态量化编码的深度神经网络压缩方法

A Dynamic Quantization Coding Based Deep Neural Network Compression Method

自动化学报. 2019, 45(10): 1960–1968 <https://doi.org/10.16383/j.aas.c180554>

基于肌电惯性融合的人体运动估计: 高斯滤波网络方法

Human Motion Estimation Based on EMG–Inertial Fusion: A Gaussian Filtering Network Approach

自动化学报. 2024, 50(5): 991–1000 <https://doi.org/10.16383/j.aas.c230581>

基于可穿戴式惯性传感器的人体运动跟踪方法综述

A Review on Wearable Inertial Sensor Based Human Motion Tracking

自动化学报. 2019, 45(8): 1439–1454 <https://doi.org/10.16383/j.aas.c180367>

未知非线性零和博弈最优跟踪的事件触发控制设计

Event-triggered Control Design for Optimal Tracking of Unknown Nonlinear Zero-sum Games

自动化学报. 2023, 49(1): 91–101 <https://doi.org/10.16383/j.aas.c220378>

行人惯性定位新动态: 基于神经网络的方法、性能与展望

李岩¹ 施忠臣² 侯燕青¹ 戚煜华¹ 谢良² 陈伟² 陈洪波¹ 闫野² 印二威²

摘要 行人惯性定位 (Inertial positioning, IP) 通过惯性测量单元 (Inertial measurement unit, IMU) 的测量序列来估计行人的位置, 近年来已成为解决室内或卫星信号遮挡环境下行人自主定位的重要手段. 然而, 传统惯性定位方法在双重积分时易受误差源影响导致漂移问题, 一定程度上限制了行人惯性定位在长时间长距离实际运动中的应用. 幸运的是, 基于神经网络 (Neural network, NN) 的方法能够仅从 IMU 历史数据中学习行人的运动模式并修正惯性测量值在积分时引起的漂移. 为此, 本文对近期基于深度神经网络 (Deep neural network, DNN) 的行人惯性定位进行全面综述. 首先对传统的惯性定位方法进行了简要介绍; 其次, 按照是否融入领域知识分别介绍了端到端 (End-to-end, ETE) 的神经惯性定位方法和融合领域知识的神经惯性定位方法的研究动态; 然后, 概述了行人惯性定位的基准数据集和评价指标, 并分析比较了其中一些代表性方法的优势和不足; 最后, 对该领域需要解决的关键难点问题进行了总结, 并探讨基于 DNN 的行人惯性定位未来所面临的关键挑战与发展趋势, 以期为后续的研究提供有益参考.

关键词 惯性测量单元, 位置跟踪, 神经网络, 行人航位推算, 自主导航, 移动设备

引用格式 李岩, 施忠臣, 侯燕青, 戚煜华, 谢良, 陈伟, 陈洪波, 闫野, 印二威. 行人惯性定位新动态: 基于神经网络的方法、性能与展望. 自动化学报, 2025, 51(2): 271–286

DOI 10.16383/j.aas.c240221 **CSTR** 32138.14.j.aas.c240221

Emerging Trends in Pedestrian Inertial Positioning: Neural Network-based Methods, Performance and Prospects

LI Yan¹ SHI Zhong-Chen² HOU Yan-Qing¹ QI Yu-Hua¹ XIE Liang²
CHEN Wei² CHEN Hong-Bo¹ YAN Ye² YIN Er-Wei²

Abstract Pedestrian inertial positioning (IP), which estimates a pedestrian's position through measurement sequences from an inertial measurement unit (IMU), has become an important solution for pedestrian autonomous positioning in indoor environments or areas with satellite signal blockages in recent years. However, traditional inertial positioning methods are prone to drift issues during double integration due to the influence of error sources, which to some extent limits the application of pedestrian inertial positioning in long-term, long-distance real-world motion. Fortunately, neural network (NN)-based methods can learn pedestrian motion patterns from historical IMU data and correct the drift caused by inertial measurement values during integration. Therefore, this paper presents a comprehensive review of recent developments in pedestrian inertial positioning based on deep neural network (DNN). First, a brief introduction to traditional inertial positioning methods is provided; Next, the latest research on end-to-end (ETE) neural inertial positioning methods and neural inertial positioning methods incorporating domain knowledge is reviewed; Following that, the benchmark datasets and evaluation metrics for pedestrian inertial positioning are summarized, and the advantages and disadvantages of some representative methods are analyzed and compared; Finally, the key challenges and difficulties that need to be addressed in this field are summarized, and the critical challenges and development trends of pedestrian inertial positioning based on DNN are discussed, aiming to provide useful references for subsequent research.

Key words Inertial measurement unit (IMU), position tracking, neural network (NN), pedestrian dead reckoning (PDR), autonomous navigation, mobile devices

Citation Li Yan, Shi Zhong-Chen, Hou Yan-Qing, Qi Yu-Hua, Xie Liang, Chen Wei, Chen Hong-Bo, Yan Ye, Yin Er-Wei. Emerging trends in pedestrian inertial positioning: Neural network-based methods, performance and prospects. *Acta Automatica Sinica*, 2025, 51(2): 271–286

收稿日期 2024-04-22 录用日期 2024-08-27

Manuscript received April 22, 2024; accepted August 27, 2024
国家自然科学基金 (62332019, 62076250), 国家重点研发计划 (2023YFF1203900, 2020YFA0713502) 资助

Supported by National Natural Science Foundation of China (62332019, 62076250) and National Key Research and Development Program of China (2023YFF1203900, 2020YFA0713502)

本文责任编辑 王雪松

Recommended by Associate Editor WANG Xue-Song

1. 中山大学系统科学与工程学院 广州 510275 2. 军事科学院
国防科技创新研究院 北京 100071

1. School of Systems Science and Engineering, Sun Yat-Sen University, Guangzhou 510275 2. Defense Innovation Institute, Academy of Military Sciences, Beijing 100071

快速而准确的自主定位对行人导航、机器人导航和工业制造自动化具有重要的作用. 尽管全球导航卫星系统 (Global navigation satellite system, GNSS) 已成为大多数定位问题的首选解决方案^[1], 但在室内环境或受到建筑物、地下通道、隧道等结构物的遮挡环境下, GNSS 信号会受到严重的阻塞和干扰, 从而无法提供连续的定位服务^[2]. 随着微机电系统 (Micro-electro-mechanical systems, MEMS) 技术的不断发展, 惯性测量单元 (Inertial measurement unit, IMU) 变得更小、更节能和更低成本. 与其他常用的传感器 (如 GNSS^[1]、无线电^[3-4]、相机^[5] 和蓝牙^[6-7]) 不同, IMU 具有自主、不受遮挡干扰等特点, 可提供连续稳定的测量值. 从 IMU 测量值中获取准确的位置估计一直是工业界和学术界的研究热点, 在定位导航、紧急服务、健康监测、军事领域和公共安全等领域中扮演着越来越重要的角色^[8-12].

惯性定位 (Inertial positioning, IP)^[13-15], 又称为惯性导航 (Inertial navigation, IN), 是基于牛顿力学的定位技术, 可以通过加速度和角速度来推断移动主体的运动状态 (速度、方向和位置). 由此, 基于惯性数据的行人定位方法^[16] 也逐渐成为一个研究热点^[17], 该方法利用移动设备 (例如, 智能手机、智能手环和 AR/VR 设备等) 中的 IMU 内置的陀螺仪和加速度计连续采样角速度和加速度, 实现了对行人自身位置的连续跟踪, 如图 1 所示.

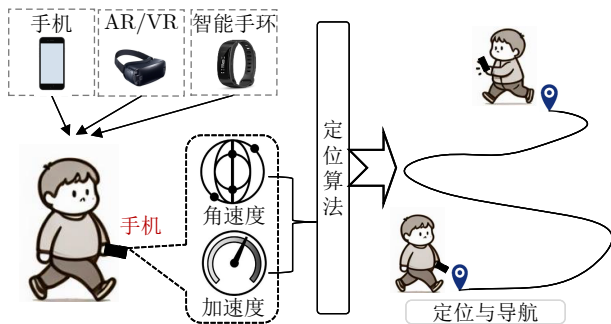


图 1 行人惯性定位范式

Fig. 1 Pedestrian inertial positioning paradigm

传统的捷联惯性导航系统 (Strapdown inertial navigation system, SINS) 算法基于牛顿运动定律和数学规则设计. 具体而言, 首先通过对角速度的积分获得姿态信息, 随后将加速度测量值转换至全局坐标系中, 再通过对转换后的加速度进行二次积分, 最终获取位置信息. 然而, 由于制造精度和工艺的限制, 低成本 MEMS IMU 的测量值常包含偏置误差、正交偏差、温度相关误差、随机噪声和随

机游走噪声等误差源^[18-19]. 在实际应用中, 这些测量误差可能会对没有约束的惯性定位系统造成显著影响, 甚至在较短时间内导致系统性能下降. 在这个过程中, 即使是微小的误差也会呈指数增长, 导致误差漂移迅速累积. 因此, 传统的捷联惯性导航系统在处理长序列惯性数据时面临着巨大挑战.

为应对上述挑战, 学者们尝试将领域知识融入惯性解算过程中来缓解行人定位过程中的误差漂移问题. 例如, 将 IMU 附着在行人脚上来检测零速相位, 并将其作为伪测量输入到卡尔曼滤波中来修正捷联惯性导航系统的状态. 另一方面, 研究人员将人类行走的周期性融入位置解算中, 利用 IMU 数据进行行人步态检测、步长估计和航向估计来实时更新位置. 然而, 此类传统方法也很难用于长时间长距离运动的真实场景. 一个制约因素在于领域局限性无法解决惯性定位的根本问题, 例如, 步长、航向和零速检测的失败也会导致系统漂移较大.

由于惯性数据存在噪声扰动以及真实环境复杂多变使得难以建立准确且通用的误差模型并投入到实际应用中. 近年来, 深度学习 (Deep neural network, DNN) 在图像、自然语言处理和语音处理等领域表现出了卓越的性能^[20-21]. 受此启发, 一些研究人员逐渐将神经网络 (Neural network, NN) 引入到行人惯性定位领域来应对上述挑战. 这种数据驱动的学习范式利用神经网络强大的非线性映射能力, 从惯性数据中提取行人运动的高级隐藏特征并回归速度和位置等状态信息, 有效地缓解了惯性传感器数据中的噪声和累积漂移问题, 显著提升了行人惯性定位的精度与稳定性^[8]. 由于此类方法在不依赖于具体的物理或数学模型的前提下展现出一定的潜在优势, 所以受到学术界和工业界的广泛关注.

为此, 本文详细地呈现了基于神经网络的行人惯性定位的研究脉络, 为了较为全面地阐述和整理相关文献, 本文重点从以下几点展开综述: 1) 简要概述传统的行人惯性定位方法及其局限性; 2) 对基于神经网络的行人惯性定位方法和现状进行详细整理; 3) 介绍具有代表性的公共数据集、评估指标以及基于神经网络方法和传统方法的性能比较; 4) 总结当前研究及应用中存在的亟待突破的关键问题, 并对可能的研究方向进行阐述与说明, 为后续研究提供借鉴. 本文结构如图 2 所示.

1 传统惯性定位系统

早期的行人惯性定位技术主要依赖于惯性测量单元中的加速度计和陀螺仪, 通过对传感器数据的

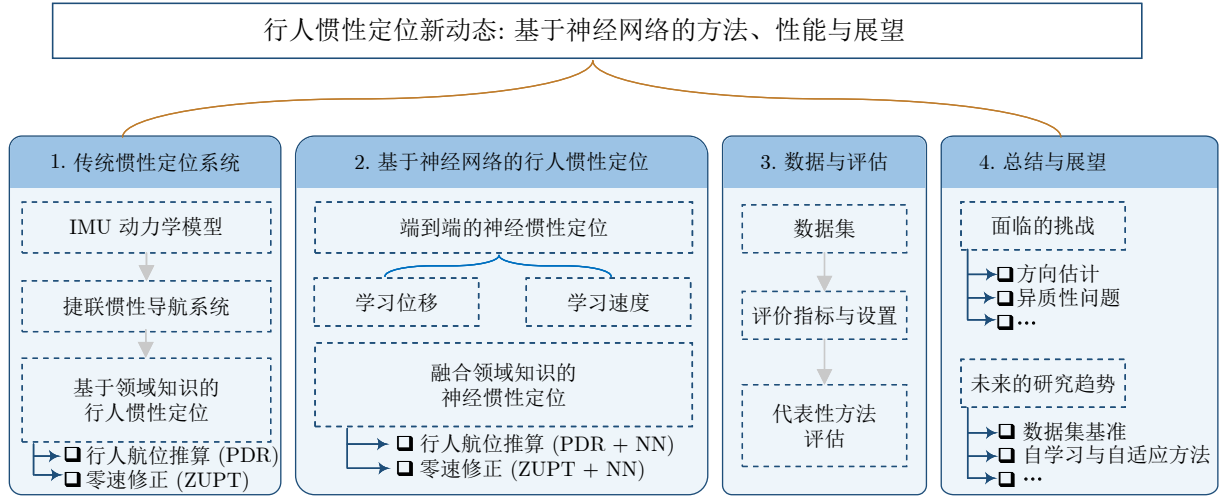


图 2 全文组织结构

Fig.2 The organization structure of this paper

积分和滤波, 实现对行人位置的估计. 传统的行人惯性定位技术经历了从捷联惯性导航到卡尔曼滤波器改进, 再到步态检测和步长估计等融入领域知识的定位方案的发展过程, 虽然定位精度不断提高, 但仍面临传感器漂移和环境变化等挑战.

本节概述了经典的行人惯性定位方法, 并总结了它们的局限性. 首先给出 IMU 动力学模型和经典的捷联惯性导航系统; 随后, 讨论两种基于领域知识减少惯性导航系统漂移的解决方案, 即行人航位推算 (Pedestrian dead reckoning, PDR) 和零速修正 (Zero velocity update, ZUPT).

1.1 IMU 动力学模型

基于惯性传感器的串行式位置解算过程不仅依赖于加速度计的测量, 还会用到陀螺仪的测量, 这将会引入很多误差源的扰动并造成误差的迅速累积. 这些误差源可以分为两类: 确定性误差和随机误差^[22-23]. 确定性误差包括轴偏、尺度误差、零偏误差、安装误差和温度误差. 随机误差包括由长期操作产生的随机传感器噪声和随机游走噪声, 这会使误差的校准和模型构建具有挑战性. 本文将这些误差对于 IMU 提供的 3D 角速度 ($\omega \in \mathbf{R}^3$) 和 3D 加速度 ($\alpha \in \mathbf{R}^3$) 测量的影响用以下公式表示:

$$\omega_t = r_t^\omega + b_t^\omega + n_t^\omega \quad (1)$$

$$\alpha_t = r_t^\alpha + b_t^\alpha + n_t^\alpha \quad (2)$$

其中, r_t^ω 和 r_t^α 分别表示在时间戳 t 陀螺仪和加速度计的真实传感器测量值; b_t^ω 和 b_t^α 代表偏差; $n_t^\omega \sim N(0, \sigma_\omega^2)$ 和 $n_t^\alpha \sim N(0, \sigma_\alpha^2)$ 分别是陀螺仪和加速度计的噪声分量.

1.2 捷联惯性导航系统

由于一些传感器信号容易受复杂环境的影响 (如遮挡、动态环境) 而退化, 这直接导致移动主体的自主定位变得困难. 基于牛顿动力学原理^[24-26] 的捷联惯性导航系统被提出来解决上述困难. 该方法利用 IMU 测量序列估计运动主体 (如, 行人) 的状态 (即位置、速度和方向), 从而大大提高了在复杂环境中实现自主定位的可能, 如图 3 所示, 整个流程可以用以下公式表示:

$$\mathbf{R}_b^n(t) = \mathbf{R}_b^n(t-1) \otimes \Omega(t) \quad (3)$$

$$\Omega(t) = \exp\left(\frac{dt}{2}\omega_{t-1}\right) \quad (4)$$

$$\mathbf{v}^n(t) = \mathbf{v}^n(t-1) + \Delta(t) \quad (5)$$

$$\Delta(t) = (\mathbf{R}_b^n(t-1) \odot \alpha_{t-1} - \mathbf{g}^n)dt \quad (6)$$

$$\mathbf{P}^n(t) = \mathbf{P}^n(t-1) + \mathbf{v}^n(t-1)dt \quad (7)$$

其中, \mathbf{R}_b^n 可用于将 IMU 测量 $\mathbf{x} \in [\omega, \alpha]$ 从本体坐标系 b 旋转到导航坐标系 n , 用表达式 $\mathbf{R}_b^n \odot \mathbf{x} = \mathbf{R}_b^n \otimes \mathbf{x} \otimes (\mathbf{R}_b^n)^T$ 表示, 其中 \otimes 是两个四元数之间的汉密尔顿乘积. t 时刻的 $\mathbf{R}_b^n(t)$ 通过在两个离散时

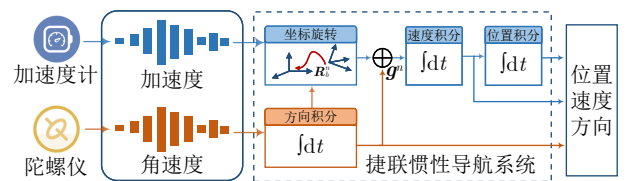


图 3 捷联惯性导航系统

Fig.3 Strapdown inertial navigation system

刻 t 和 $t-1$ 之间的相对方向 $\Omega(t)$ 进行更新, 如式 (3) 和 (4) 所示. 为了跟踪速度 $v^n(t)$ 和位置 $P^n(t)$, 使用 $R_b^n(t-1)$ 将 $\alpha(t-1)$ 旋转到导航坐标系中并消除重力加速度 g^n 的影响, 然后通过式 (6) 和 (7) 来迭代更新速度和位置.

捷联惯性导航系统虽然方法简单, 但是 IMU 的测量值会受到误差源的影响, 在长时间长距离实际应用中, 出现显著的漂移问题. 在 5 ~ 10 min 的行人运动场景中, 捷联惯性导航系统的绝对轨迹误差达到了上百米^[27]. 为了减小误差, 有学者引入了卡尔曼滤波器等方法, 通过对传感器数据进行融合和滤波, 提高了定位精度. 然而, 这些方法在处理非线性和非高斯噪声时仍然表现不佳.

1.3 基于领域知识的行人惯性定位

先前的方法由于传感器固有的特性, 会导致漂移和定位困难等问题. 为了解决上述问题, 研究人员通过观察行人运动规律的周期特征, 提出引入领域知识来增加约束, 进而实现漂移校正和精确的行人定位, 最具代表性的方法有行人航位推算^[28-29]和零速修正^[30-31].

1.3.1 行人航位推算

基于行人航位推算 (PDR) 的惯性定位通过检测行人步长与航向的策略来更新位置. 这种策略有助于防止误差不断累积, 从而减轻轨迹漂移问题. PDR 的推算过程^[32]如图 4 所示, 主要包括四个步骤: 步态检测、步长估计、航向估计和位置更新. 在步态检测过程中, PDR 基于阈值和峰值检测^[33]、相关性检测^[34]和谱分析^[35]等技术来识别步态并分割相应的惯性数据; 然后, 通过 Weinberg 公式^[36]来估计步长, 同时积分陀螺仪信号来估计航向; 最后, 利用估计的航向和步长来更新行人的位置. 通过避免加速度的双重积分并结合可靠的步长估计模型, PDR 有效地减少了惯性定位漂移. 然而, PDR 方法通常依赖于人类行走的周期性特征, 仍然可能发生不准确的步态检测和步长估计, 导致系统误差漂移较大. 在 5 ~ 10 min 的行人运动场景中, 行人航位推算的绝对轨迹误差达到了 30 m 左右^[27].

1.3.2 零速修正

人类行走除了周期性之外, 还存在一个普遍规律, 即当一只脚完全着地会存在暂时静止的那一时刻. 基于此, 研究人员通过识别和利用这些短暂的静止时刻, 可以有效地校正惯性传感器数据中的累积误差, 从而提高定位精度. 整个算法流程如图 5 所示, 零速修正 (ZUPT) 算法旨在通过识别人类行走的

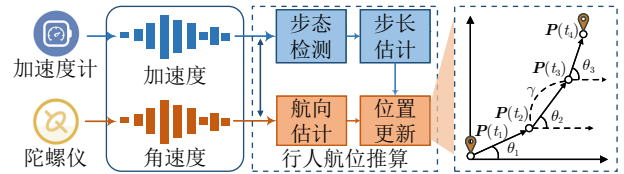


图 4 行人航位推算

Fig. 4 Pedestrian dead reckoning

静止阶段并使用零速度作为卡尔曼滤波器^[30]中的观测来补偿 SINS 的误差^[37]. 为了便于检测静止阶段, 首先将惯性传感器与脚固连, 然后利用峰值检测^[38]、过零检测^[39]或自相关^[40]等技术来分析惯性数据并分割零速静止相位. 一旦检测到静止相位, 就将零速作为伪观测量输入到卡尔曼滤波过程中, 进而输出可靠的位置信息. 然而, ZUPT 的有效性取决于用户脚仍然完全静止以及 IMU 附着在用户脚上的假设, 任何不正确的静止相位检测或小运动干扰都会导致定位系统漂移, 安装于脚上的方式也限制了其应用场景, 如当手机放置于口袋时则无法使用此类方法. 此外, 这些方法容易受到传感器噪声和运动模式的影响, 进而限制了零速度区间的检测效果^[41].

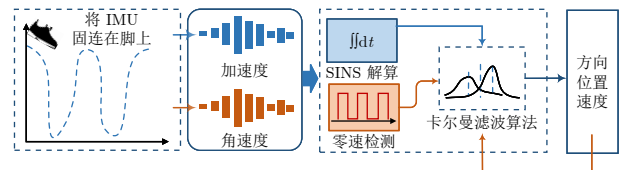


图 5 零速修正

Fig. 5 Zero velocity update

基于传统方法的零速修正方法在 10 min 自然运动中的绝对轨迹误差具体数值会因不同实验条件和传感器设备而有所不同. 不过, 一般情况下, 这种误差通常在几十米到几百米之间.

2 基于神经网络的行人惯性定位

传统的行人惯性定位方法基于运动学模型与人类行走规律, 受限于噪声和漂移的影响, 想要实现复杂动态环境下的高精度自主定位是困难的. 近年来, 神经网络为行人惯性定位的研究提供了新的思路, 它可以学习潜在的高级特征来校正漂移并实现动态环境中的定位. 此部分根据是否引入领域知识来约束惯性定位, 将基于神经网络的行人惯性定位方法分为端到端 (End-to-end, ETE) 的神经惯性定位方法和融合领域知识的神经惯性定位方法. 本节将对所有代表性方法做详细分析与介绍.

2.1 端到端的神经惯性定位

传统的捷联惯性导航系统通常面临三个主要问题: 一是惯性数据容易受到噪声影响, 导致误差漂移问题; 二是传统的定位方法通常依赖于动力学模型, 无法有效地捕获惯性数据中复杂的非线性关系; 三是传统算法需要定期的人为校准, 行人运动模式(如, 走、跑、静止)的复杂性和移动设备携带方式(如, 手、裤兜、书包)的多变性可能会提高位置解算的复杂度. 为解决上述问题, 研究人员开始探索使用端到端的神经惯性定位, 利用神经网络强大的非线性映射能力以缓解定位漂移并学习行人复杂的运动模式. 这些模型可以部分或完全取代传统的捷联惯性导航系统, 仅从 IMU 传感器测量序列估计行人的位置信息. 如图 6 展示了端到端的神经惯性定位的一般范式. 现有的神经惯性定位基于学习目标主要分为学习位移和学习速度两类^[42], 图 7 展示了它的大致流程, 本节将对其展开详细讨论.

2.1.1 学习位移

位移是定位系统中最直观且不可或缺的状态信息之一. 给定行人的一个初始状态(位置和方向), 就可以通过计算某一固定时间段下的位移来实时更新行人的位置信息. 然而, 惯性传感器受到偏差和噪声的困扰, 当加速度被双积分以获得位移时, 误差会持续累积. 因此, Chen 等^[43]首次提出 IONet 避免连续积分的循环, 将惯性定位问题转换为序列学习问题, 该方法首先将 IMU 原始数据序列分割成独立的窗口, 然后利用长短时记忆网络(Long short-term memory, LSTM)^[44]的端到端学习框架将这些独立窗口映射成极坐标系下的相对位移, 最后通过将运动位移与初始位置相加来重建轨迹. 另

外, 鲁棒的神经网络模型需要大量的数据来驱动, 据此, 该文基于行人以不同方式携带智能手机在高精度视觉运动跟踪系统的房间中行走收集数据, 智能手机记录 IMU 数据, 高精度的运动捕捉系统捕捉运动真值. 在 5 ~ 10 min 的行人运动跟踪场景中, IONet 的绝对轨迹误差稳定在 20 m 左右, 优于传统的 PDR 算法.

尽管 IONet 展现出了一定的潜力, 但它较高的计算和内存需求会限制深度神经网络模型部署到移动端设备^[45]. 基于此, 文献 [8] 设计了一个高效、快速的 DNN 模型, 称为 L-IONet, 它是 IONet 的一个扩展变体. 它利用自回归模型 WaveNet^[46] 替换双向 LSTM (Bidirectional LSTM, BiLSTM) 来处理惯性传感器的长连续信号并预测极坐标系下的相对位移, 舍弃了循环计算和复杂的非线性操作并使用并行计算的方式提高了模型的推理速度. EqNIO^[47] 引入 IMU 数据中固有的物理旋转反射对称性, 进而提高了模型的泛化性.

先前的方法主要侧重于从多设备携带方式的数据中获得紧凑的惯性表征, 然而却忽略了行人运动模式(如, 走、跑、静止)的复杂性. 因此, 在单一运动域特定数据集上训练的模型可能不能很好地泛化到新的运动域. 针对这一问题, 文献 [48] 提出了 MotionTransformer, 通过生成对抗网络(Generative adversarial network, GAN)^[49]和域适应^[50]来学习一个领域不变的语义表示, 让惯性定位模型对新的运动领域进行自适应.

Liu 等^[51]将 2D 定位问题扩展成 3D 状态估计问题, 首先基于卷积神经网络(Convolutional neural network, CNN)回归两个时刻之间的 3D 相对位移和不确定性, 然后将二者合并到扩展卡尔曼滤波器

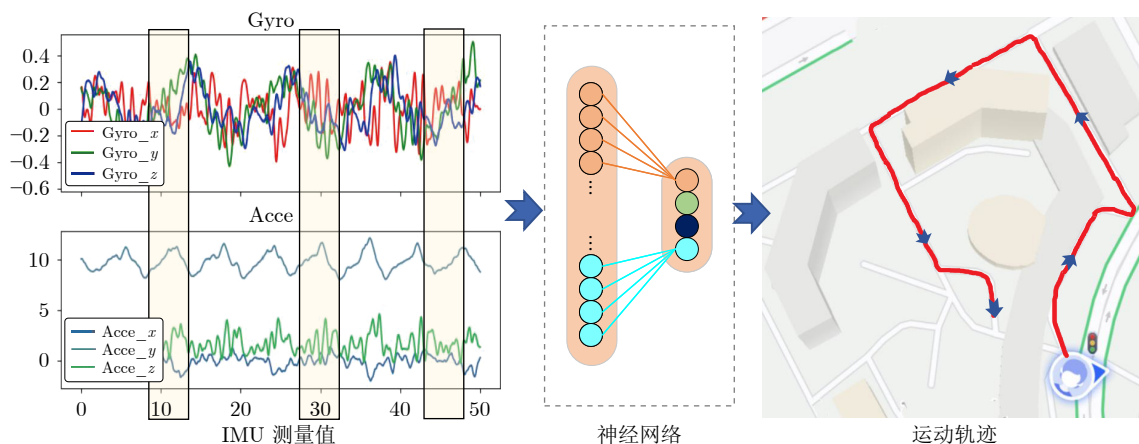


图 6 基于神经网络的行人惯性定位范式

Fig.6 Paradigm of pedestrian inertial positioning based on neural network

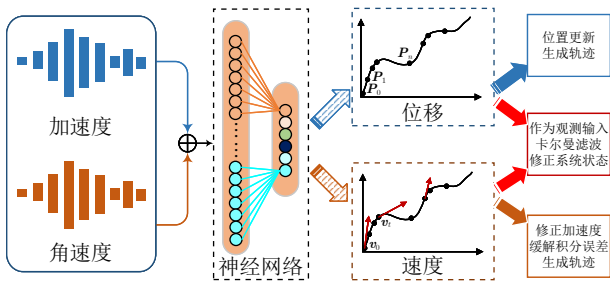


图7 神经惯性定位算法流程图

Fig.7 Neural inertial positioning algorithm flowchart

(Extended Kalman filter, EKF)^[52] 中作为观测进行状态估计, 包括方向、速度、位置和 IMU 偏差。

2.1.2 学习速度

基于学习速度的行人惯性定位是指将加速度和角速度作为一个整体进行编码从而得到一个潜在的特征向量. 这个特征向量能够捕捉到行人运动的高级特征和不同信号之间的非线性交互关系, 并且可用于下一步回归速度信息. 这种学习的过程避免了速度积分的循环, 可以在一定程度上修正加速度并校正漂移. 其中, 最经典的一个例子是 Herath 等^[27] 提出的 RoNIN, 它首先通过训练一个深度神经网络 (包括基于 ResNet、LSTM 和 TCN (Temporal convolutional network) 的三种变体) 从惯性数据中预测行人的 2D 速度向量, 然后将学习到的速度向量和惯性数据转换到 Z 轴与重力对齐的坐标系中, 最后对速度进行积分以估计位置. 该方法设计了三种基线模型, 分别从时间和空间维度提取高级特征来回归速度, 在 5 ~ 10 min 的行人运动跟踪场景中, RoNIN 三种变体的绝对轨迹误差均稳定在 5 m 左右, 优于传统的 PDR 算法. 由于 IMU 原始数据序列分割成独立的窗口, RoNIN-ResNet 的网络结构可以捕获到更加丰富的上下文信息, 后续很多工作扩展 RoNIN-ResNet 并做出相应的改进. 例如, Wang 等^[53] 改变了 RoNIN 的回归目标, 通过训练 ResNet 网络来回归速度大小和移动角度, 基于此可以实时获取行人的运动航向. 尽管 RoNIN-ResNet 展现出一定的潜力, 但由于参数过多, 将其部署到移动设备上困难的. 为此, Zeinali 等^[54] 通过结合深度和点卷积扩展了 RoNIN-ResNet, 该方法在准确性和效率方面均有所提升, 并且能够部署在移动设备上. 大量数据的高质量标注是基于神经网络的惯性定位面临的一大难题, 为减小神经网络对数据的依赖, Cao 等^[55] 以 RoNIN-ResNet 为基线模型, 引入了旋转等方差作为自监督信号来训练惯性定位模型. 该模型减少了对大量标记数据进行训练的依赖, 并使得利用各种未标记数据来更新模型成为

可能. ResMixer^[56] 是一种轻量化的网络结构, 它用 Mixer 层替换了深度卷积层, 不仅捕获输入数据不同部分之间的交互或依赖关系, 同时可以更好地实现移动端部署.

先前的基于 ResNet 的方法主要侧重于扩展原始测量以获得高维的惯性表征, 然而却忽略了同时捕捉不同时间和不同模态输入信号之间的复杂相互依赖性. 为了更好地从单个模态惯性信号中学习时间依赖性以及每个时间戳中所有信号之间的相关性, 研究人员开始尝试在神经惯性定位中引入各种注意力机制. Wang 等^[57] 提出一种用于行人惯性导航的时间注意混合神经网络, 利用 CNN 提取空间特征, 利用 LSTM 捕获全局时间信息, 并利用时间注意力机制对 LSTM 产生的隐藏状态进行加权. Zhu 等^[58] 首先将归一化注意力模块 (Normalization-based attention module, NAM) 和卷积块注意力模块 (Convolutional block attention module, CB-AM) 添加到残差网络 ResNet 中, 以增强网络对通道和空间特征的学习能力, 然后添加 BiLSTM 模块可以增强网络在长距离学习的能力. 与 RoNIN-ResNet 相比, 该方法的平均均方根误差 (Root mean square error, RMSE) 值降低了 1.495 m, 证明了该方法有效地提高了行人定位的精度. Chen 等^[59] 则在速度回归网络中融入了 Res2Net^[60] 模块, 用来提取更加细粒度的特征表示, 让网络关注更有价值的特征线索, 以此来提高网络预测的质量.

最近, 自注意力机制^[61-62] 因其具有强大的表征能力、时序建模能力与可扩展性进一步提升了深度神经网络的学习能力, 尤其是在处理文本、音频等时序信号时尤为突出. 基于此, 很多研究工作开始利用自注意力机制来探索 IMU 测量在空间和时间维度上的丰富上下文信息以用于惯性特征表示. Rao 等^[26] 提出一种鲁棒的基于自注意力的模型, 用于从 IMU 测量序列中回归 2D 速度和轨迹. 该模型首先利用局部和全局自我注意层扩展 RoNIN-ResNet, 通过高效注意矩阵的学习来捕获空间上下文信息并增强惯性表征能力. 其次, 通过利用 Transformer^[61] 解码器中的多头注意力, 进一步将这些空间表示与时间知识融合. 最后, 利用不确定性的多任务学习来提高 2D 速度的学习效率和预测精度. 先前的工作仅仅可以输出相对位置, 想要得到绝对定位还需从起始位置进行累积计算. 为此, Herath 等^[63] 指出人体的运动模式在不同时间和不同位置具有独特性, 可以将其作为一种被识别信号来确定绝对位置. 针对这一想法, 该文献首先从惯性数据计算一系列速度, 然后使用基于 Transformer 的 DNN 框架将速

度序列转换为绝对位置。

尽管当前神经惯性定位的工作已经取得一定的进展, 但是这些方法高度依赖于设备方向将惯性数据旋转到全局坐标系中. 为了解决这个问题, Sun 等^[64] 提出一个两阶段位置估计过程, 它首先借助磁力计和陀螺仪的测量数据来估计设备方向, 通过最小化对设备方向的依赖将惯性数据旋转到全局坐标系中, 然后这些数据被传递到基于 BiLSTM 的位置模块, 该模块经过训练以最小化每个窗口的位移误差来进行定位. 实验结果表明, 与基于 ResNet 的方法相比, IDOL 可以获得更准确的定位精度. 除此之外, Wang 等^[65] 基于 DNN 来推断本体坐标系中的行走速度, 并将其作为观测与不变扩展卡尔曼滤波 (Invariant extended Kalman filter, IEKF) 相结合来降低线性加速度中的噪声.

此类算法在 5 ~ 10 min 的基于智能手机的行人自然运动跟踪场景中, 绝对轨迹误差均稳定在 3 ~ 6 m 之间. 然而, 此类算法也存在一些挑战: 如何设计合适的惯性表征是一个关键问题, 过于简单可能无法充分利用惯性序列的非线性交互信息, 过于复杂则会增加计算复杂度并导致过拟合. 因此, 未来需要进一步提高算法的性能与应用范围.

2.2 融合领域知识的神经惯性定位

目前传统的融合领域知识的行人惯性定位面临两大难题: 行人都是不同的个体, 各种特征 (包括身高、腿长、年龄、性别、体重以及个体的步行习惯) 均会导致步长在不同行人之间存在差异, 进而导致定位表现可能不稳定; 不正确的步态、步长、航向、静止相位检测或小运动干扰都会导致定位系统漂移, 这使得这种方法不适合日常使用. 为了解决上述问题, 研究学者就如何利用领域知识和神经网络来优化行人惯性定位也进行了不少尝试, 包括行人航位推算 (PDR + NN) 和零速修正 (ZUPT + NN) 两种方式. 图 8 和图 9 分别展示了两种方式的算法流程, 本节将对以上内容展开详细介绍.

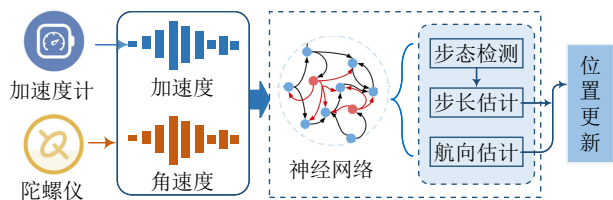


图 8 PDR + NN 流程图

Fig.8 Flowchart of PDR + NN

2.2.1 PDR + NN

早期的行人航位推算利用惯性传感器对步数进

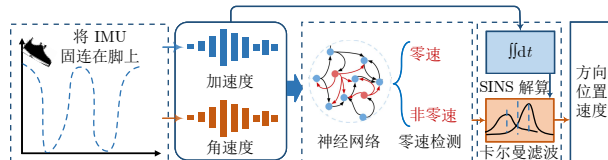


图 9 ZUPT + NN 流程图

Fig.9 Flowchart of ZUPT + NN

行计数, 并估计步长, 结合航向输出来推算当前位置信息. 行人个体间的异质性导致步态以及运动模式的多样性, 会引发行人航位推算过程中的估计误差, 因此算法存在误差漂移以及适应与泛化能力不足等问题. 为了解决这些问题, 研究人员将神经网络技术纳入步态检测、动态步长估计和步行航向估计的过程中, 能够更有效地校正漂移、提高鲁棒性和适应性.

1) 步态检测可以准确地识别行人行走过程中每一步的发生, 并据此计算步数. 行人步态因个体差异、运动模式和环境变化而多样且复杂, 传统方法难以全面准确地识别不同步态. 为了解决上述问题, 最基本的步态计数器都是基于 CNN^[66] 和 LSTM^[67] 来提取行人步态中的一致性特征, 并为后续推算位置提供重要指导. 这类步态计数器在大多数情况下都能稳健地工作, 但是在停止行走时模型的输出表现出衰减的振荡现象. 另外, 有时当一个人开始走路时可能会遗漏前两个或几个步数. 为此, Han 等^[68] 提出一个端到端的行人航位推算框架, 称为 WAIT, 它是一个基于 ResNet 的自动编码器结构, 可以将被误差破坏的原始 IMU 测量值转化为无误差的波形, 从生成的波形中可以提取各种与移动性相关的信息, 不仅包括单个位置, 还包括步数、航向变化、速度和完整的行人轨迹等; 而且, 其性能不会因为智能手机携带模式或用户数量不同而显著下降. 通过三种不同智能手机携带方式和两个用户进行现场实验, 结果表明, 与传统的 PDR 算法相比, 平均定位误差下降了 0.5 m 左右.

2) 步长估计是 PDR 的关键组成部分, 对 PDR 的性能有重要的影响. 早期的步长估计模型存在一些局限性, 例如需要了解行人的身高, 缺乏考虑不同的手机携带方式以及依赖于空间约束. 为了解决这些问题, Gu 等^[69] 提出一种基于堆叠的自动编码器^[70] 的步长估计模型, 主要包括分割、特征学习和步长估计模块, 该模型可以适应不同的手机携带方式, 并且不需要个人身高信息和空间约束. StepNet^[71] 是一种基于 CNN 的 PDR 网络, 它可以动态地回归步长或距离的变化. 与传统的静态步长估计相比, 它实现了显著的性能提升, 错误率仅为 2.1% ~ 3.2%.

上述的步长估计算法在以正常速度行走直线的情况下表现相对较好,但在复杂场景中其误差会过度增大.不准确的行走距离估计导致行人航位推算的累积定位误差较大.为此,Wang等^[72]提出一种基于LSTM网络和去噪自动编码器(Denoising autoencoder, DAE)的步长估计方法.其中,LSTM网络用于捕获时间依赖关系,并从损坏的惯性传感器测量中提取重要的关键特征;然后,采用DAE自动消除关键特征中的固有噪声,得到去噪的特征向量;最后,回归模块将去噪的特征向量映射得到步长.大量的实验结果表明,该方法优于其他先进的算法,单次步长错误率为4.63%,步行距离错误率为1.43%.

然而,以上方法没有考虑日常使用中智能手机携带方式(即传感器位置)的变化往往会导致较大的估计误差,为了适配实际应用场景,Bo等^[73]提出一种新的步长估计框架,称为模式无关神经网络(Mode-independent neural network, MINN).首先,提出一种基于多级ResNet和门控循环单元(Gated recurrent unit, GRU)的分层神经网络时空特征提取方法;然后,使用对抗训练和子类分类器来构建一个多源无监督域适应网络,该网络可以从不同模式中提取数据共享的模式不变特征.

3) 航向估计可以实时提供行人移动方向信息.但是陀螺仪漂移和噪声会导致航向估计中误差的累积.因此有学者利用神经网络(例如,LSTM)来对陀螺仪进行矫正、补偿漂移^[74-75].Wang等^[76]利用空间Transformer网络^[77]来对齐本体坐标系和导航坐标系,并基于分层LSTM的序列到序列模型准确估计行人的航向.

然而,设备航向与行人航向之间的错位使得根据传感器数据估计真实的步行航向是困难的.为了解决这个问题,Manos等^[78]利用时间卷积和多尺度注意层提取运动矢量进行航向估计,该网络模型与几何计算相结合,利用重力和地磁方向,将运动矢量转换为相对于北的航向角.该方法得到的航向误差中值为 9.8° ,与基线方法相比降低了 2.5° .PDRNet^[79]基于ResNet设计了一个用于智能手机位置识别的分类网络和一个获取距离和航向变化的回归网络.它遵循传统PDR算法的过程,但用神经网络代替了步长和航向估计模块.在RIDI数据集上的实验结果表明,PDRNet与其他方法相比取得了最先进的性能,比传统PDR和RoNIN的绝对轨迹误差分别提高了67%以上和30%以上.

2.2.2 ZUPT + NN

基于零速修正的传统行人惯性定位系统一般会

通过检测行人短时间内停止移动的零速度时刻来校正惯性定位系统的误差.虽然ZUPT可以提高惯性导航系统的精度,但是先前的零速识别依赖于手工设计的规则或简单的阈值方法,易受传感器误差、行人动作变化或环境干扰的影响,进而导致误检测或漏检.为了解决这个问题,研究人员探索了利用神经网络模型学习从原始传感器数据中直接学习到步态特征和行人行走模式,利用深度学习强大的特征提取和分类能力对用户是否在ZUPT阶段进行分类,以此提高零速修正的可靠性和精度.为了解决上述问题,Wagstaff等^[80]提出一种用LSTM代替零速度检测器来提高零速度辅助惯性定位系统精度的方法.LSTM输入一系列IMU数据,通常为100个连续的数据点,并输出用户在当前时间步长静止或运动的概率;然后将基于LSTM的零速度检测的结果输入到扩展卡尔曼滤波器中.与现有的固定阈值零速度检测器相比,3D定位误差降低了34%以上.Yu等^[81]提出一种基于卷积神经网络的零速度点探测器方法,它可以自适应地选择不同行人(如步行、快速步行和跑步)中的零速度点,该方法相比于固定阈值算法表现得更加鲁棒.

表1中提供了基于神经网络的行人惯性定位现有的最具代表性工作的概览.总体来讲,基于神经网络的行人惯性定位方法能够自动提取复杂的非线性特征并适应不同个体的运动模式和行为差异,能有效校正传感器漂移和误差,提高长期定位精度.但是此类方法也存在一些劣势,包括对大量标注数据的需求较高,数据收集和标注成本较高;计算资源需求较高,特别是在实时定位应用中需要高性能的处理器和较大的存储空间;模型设计和调参过程复杂,需要较高的专业知识和经验.

3 数据与评估

除了建立网络模型外,模型的训练、优化、评估也至关重要.本节将对基于神经网络的行人惯性定位现有的公开数据集与评估方法进行总结,同时比较部分主流方法的性能来衡量模型的泛化能力.

3.1 数据集

基于神经网络的行人惯性定位需要使用海量的数据来提高泛化能力和导航定位的精度.现有的文献所涉及的一般公共数据集大多都是以智能手机为数据采集平台,主要有RIDI^[82]、TUM-VI^[83]、OXIOD^[84]、RoNIN^[27]、IDOL^[64]、CTIN^[26]、SIMD^[85].表2详细描述了所有数据集的基本属性,其中,RIDI^[82]、OXIOD^[84]、RoNIN^[27]是比较常见的公开数据集.

● RIDI¹ 是一个用于行人惯性定位的基准数据集, 该数据集由 10 名使用具有 3D 跟踪功能的智能手机的被试者在 2.5 小时内收集完成. 其中 IMU 数

据由智能手机采用不同的携带方式所采集 (包括手持、包、裤袋和胸前). 该数据集的真值轨迹由 Tango 手机内置的视觉惯性里程计² 来记录.

表 1 基于神经网络的行人惯性定位方法概览
Table 1 Overview of neural network-based pedestrian inertial positioning methods

方法	年份	模型	学习方式	方法特征
IONet ^[48]	2018	LSTM	监督	将惯性定位问题转换为序列学习问题, 基于 LSTM 来学习位移并构造惯性里程计
L-IONet ^[6]	2020	WaveNet	监督	利用自回归模型替换 LSTM 来处理长序列惯性信号并预测极坐标系下的位移
Motiontransformer ^[48]	2019	LSTM	监督	通过生成对抗网络和域适应来学习一个领域不变的语义表示
TLIO ^[54]	2020	CNN	监督	基于 CNN 回归相对位移和不确定性并将二者合并到卡尔曼滤波器进行状态估计
RoNIN ^[27]	2020	CNN/LSTM	监督	基于 CNN/LSTM 从惯性数据中预测行人的 2D 速度向量
Wang 等 ^[53]	2021	CNN	监督	通过 ResNet 来回归速度大小和移动角度
IMUNet ^[54]	2024	CNN	监督	使用深度和点卷积替换传统卷积操作提高模型推理速度
RIO ^[55]	2022	CNN	自监督	引入旋转等方差作为强大的自监督信号来训练惯性定位模型
HNNNTA ^[57]	2022	CNN/LSTM	监督	利用时间注意力机制对 LSTM 产生的隐藏状态进行加权
RBCN ^[58]	2023	CNN/LSTM	监督	利用多种混合注意力机制增强网络对通道和空间特征的学习能力
Res2Net ^[59]	2022	CNN	监督	融入 Res2Net 模块来提取更加细粒度的特征表示
CTIN ^[26]	2022	Transformer	监督	首个基于 Transformer 来融合空间表示与时间知识的模型
RIOT ^[62]	2023	Transformer	监督	通过结合真实位置先验递归的学习运动特征和系统误差偏差
NILOC ^[63]	2022	Transformer	监督	将独一无二的人体运动模式映射成行人位置
IDOL ^[64]	2021	LSTM	监督	将行人惯性定位分为方向估计和位置估计两个阶段
Shao 等 ^[66]	2018	CNN	监督	基于深度卷积神经网络的步长检测方案, 以提高计步器的鲁棒性
Ren 等 ^[67]	2021	LSTM	监督	设计一种基于 LSTM 的步态计数器
WAIT ^[68]	2023	CNN	监督	利用自动编码器将 IMU 测量值转化为无误差的波形并提取各种与移动性相关的信息
Gu 等 ^[69]	2018	Autoencoder	监督	基于堆叠的自动编码器的步长估计模型
StepNet ^[71]	2020	CNN	监督	基于 CNN 动态的回归步长或距离的变化
Wang 等 ^[72]	2019	LSTM	监督	在步长估计模型中加入变分自编码自动消除特征向量中的固有噪声
Manos 等 ^[78]	2022	CNN	监督	利用时间卷积和多尺度注意层提取运动矢量进行航向估计
PDRNet ^[79]	2022	CNN	监督	基于 ResNet 设计一个位置识别和一个获取距离和航向变化的回归网络
Wagstaff 等 ^[80]	2018	LSTM	监督	用 LSTM 代替标准零速度检测器来辅助惯性导航系统
Yu 等 ^[81]	2019	CNN	监督	一种基于卷积神经网络的零速度点探测器
Bo 等 ^[73]	2022	ResNet/GRU	无监督	利用对抗训练和子类分类器来构建一个多源无监督域适应网络

注: 上述方法根据是否融入领域知识分为两类.

表 2 行人惯性定位数据集
Table 2 Pedestrian inertial positioning datasets

数据集	年份	采样频率	IMU 载体	真值	数据集大小 (轨迹数)	设备携带方式
RIDI ^[82]	2017	200 Hz	Lenovo Phab2 Pro	Tango 手机	74	裤袋、包、手持、胸前
TUM VI ^[83]	2018	200 Hz	—	动作捕捉系统	28	手持
OXIOD ^[84]	2018	100 Hz	iPhone 5/6/7 Plus, Nexus 5	动作捕捉系统	158	手持、口袋、手袋、推车
RoNIN ^[27]	2019	200 Hz	Galaxy S9, Pixel 2 XL	AR 设备	276	自然携带
IDOL ^[64]	2020	100 Hz	iPhone 8	Kaarta Stencil	84	自然携带
CTIN ^[26]	2021	200 Hz	Samsung Note, Galaxy	Google ARCore	100	自然携带
SIMD ^[85]	2023	50 Hz	多种型号智能手机	GPS/IMU	4562	自然携带

¹ https://www.dropbox.com/s/9zza3h3u4bta23/ridi_data_publish_v2.zip?dl=0

² <https://github.com/higerra/TangoIMURecorder>

● OXIOD³ 是学术文献中最常用的行人导航数据集, 数据量庞大且真值轨迹准确性高. 主要是由被试者以不同方式携带智能手机 (主要是 iPhone 7 Plus), 执行慢走、常规行走和奔跑等不同运动模式收集而来. 该数据集由 158 个轨迹序列组成, 涵盖超过 42 公里的距离. 真值轨迹由具有 0.5 mm 精度的高精度 Vicon 运动捕捉系统实时记录.

● RoNIN⁴ 是惯性导航领域广泛认可的基准数据集, 该数据集形式丰富, 近乎完美地复刻了人类日常生活中的自然运动. 它包含来自 100 名行人被试者超过 40 小时的 IMU 传感器数据, 并捕捉到自然人类运动中的 3D 真实轨迹. 该数据集以 200 Hz 的频率提供了 276 个序列的方向和位置信息. 数据收集协议涉及两个设备, 其中一个手机固定在身体上用于 3D 跟踪, 使被试者可以自由操纵另一个手机进行 IMU 数据收集. 此外, 该数据集目前只有一半是公开可用的.

3.2 评价指标与设置

为了评估基于神经网络的惯性定位的有效性, 研究人员通常采用两种评价指标: 绝对轨迹误差 (Absolute trajectory error, ATE) 和相对轨迹误差 (Relative trajectory error, RTE).

● ATE: 绝对轨迹误差表示估计轨迹与地面真实轨迹之间的整个轨迹上的 RMSE. ATE 充当了一种全局的度量标准, 为整个轨迹范围内的导航坐标系内的位置估计一致性提供了度量标准.

$$\text{ATE} = \sqrt{\frac{1}{m} \sum_{t=1}^m \| (x_t - \tilde{x}_t) \|^2} \quad (8)$$

其中, m 表示整个轨迹的序列长度; x_t 和 \tilde{x}_t 分别表示在 t 时刻的估计值和地面真实值.

● RTE: 一种用于评估估计轨迹在固定时间间隔或固定长度的局部准确性的基准指标. 一般计算在指定时间间隔 Δt 内 (大多设为 1 min) 估计值与地面真实值之间的平均均方根误差.

$$\text{RTE} = \sqrt{\frac{1}{m} \sum_{t=1}^m \| (x_{t+\Delta t} - x_t) - (\tilde{x}_{t+\Delta t} - \tilde{x}_t) \|^2} \quad (9)$$

目前基于神经网络的行人惯性定位模型大多使用 PyTorch 实现, 并在 NVIDIA RTX GPU 上使用 Adam 优化器进行训练. 为了满足行人日常生活, 当前使用的网络结构参数量集中在 $10^5 \sim 10^7$ 数量

级, 可以满足人类使用智能手机实现实时定位的基本要求.

3.3 代表性方法评估

表 3、表 4 和表 5 分别展示了一些代表性方法在 RIDI 数据集、OXIOD 数据集和 RoNIN 数据集上的性能结果. 其中, 经典的传统方法包括 SINS^[25]、PDR^[29] 以及两种代表性的基于学习的模型, 包括 RIDI^[82]、RoNIN^[27].

在所有的数据集上可以观察到一个一致的结果, 传统的捷联惯性导航系统明显落后于数据驱动

表 3 在 RIDI 测试数据集上的行人惯性定位方法对比 (m)

Table 3 Comparison of pedestrian inertial positioning methods on the RIDI test dataset (m)

模型	seen-ATE	seen-RTE	unseen-ATE	unseen-RTE
SINS ^[25]	31.06	37.53	32.01	38.04
PDR ^[29]	3.52	4.56	1.94	1.81
RIDI ^[82]	1.88	2.38	1.71	1.79
R-LSTM ^[27]	2.00	2.64	2.08	2.10
R-ResNet ^[27]	1.63	1.91	1.67	1.62
R-TCN ^[27]	1.66	2.16	1.66	2.26

表 4 在 OXIOD 测试数据集上的行人惯性定位方法对比 (m)

Table 4 Comparison of pedestrian inertial positioning methods on the OXIOD test dataset (m)

模型	seen-ATE	seen-RTE	unseen-ATE	unseen-RTE
SINS ^[25]	716.31	606.75	1941.41	848.55
PDR ^[29]	2.12	2.11	3.26	2.32
RIDI ^[82]	4.12	3.45	4.50	2.70
R-LSTM ^[27]	2.02	2.33	7.12	5.42
R-ResNet ^[27]	2.40	1.77	6.71	3.04
R-TCN ^[27]	2.26	2.63	7.76	5.78

表 5 在 RoNIN 测试数据集上的行人惯性定位方法对比 (m)

Table 5 Comparison of pedestrian inertial positioning methods on the RoNIN test dataset (m)

模型	seen-ATE	seen-RTE	unseen-ATE	unseen-RTE
SINS ^[25]	675.21	169.48	458.06	117.06
PDR ^[29]	29.54	21.36	27.67	23.17
RIDI ^[82]	17.06	17.50	15.66	18.91
R-LSTM ^[27]	4.18	2.63	5.32	3.58
R-ResNet ^[27]	3.54	2.67	5.14	4.37
R-TCN ^[27]	4.38	2.90	5.70	4.07

注: 其中, “seen”表示测试集和训练集的被试相同; “unseen”表示测试集中的被试在训练集中未出现过; “ATE”表示绝对轨迹误差; “RTE”表示相对轨迹误差.

³ <http://deepio.cs.ox.ac.uk/>

⁴ <https://ronin.cs.sfu.ca/>

的惯性导航系统. 这种差距主要归因于捷联惯性导航系统固有的误差传播, 这对它们的定位能力产生了不利影响. 传统的捷联惯性导航系统基于牛顿力学将 IMU 测量直接积分到位置、速度和方向上. 然而, 由于测量噪声的误差传播, SINS 快速漂移, 无法提供合理的结果. 类似地, PDR 也是行人惯性定位主流的传统解决方案, 但是不准确的步长和航向估计也会引起误差传播, 导致行人运动跟踪的失败. 此外, 基于 ResNet 架构的模型性能始终超过了使用 LSTM 和 TCN 的模型结构, 表明了基于 ResNet 的模型从 IMU 测量中提取语义信息的强大能力, 突出了基于学习的运动建模方式在复杂运动场景中的有效性.

由以上结果可知, 随着深度学习技术的融入, 行人惯性定位领域的发展变得更加多元化; 不仅仅在各项评价指标上有显著提升, 更重要的是细分的研究热点可以在本领域更好地拓展和研究.

4 总结与展望

近年来, 基于神经网络的行人惯性定位已经逐渐成为导航定位领域中新兴的研究方向, 一经提出就受到工业界和学术界的广泛关注. 随着研究的不断深入, 此类方法已经为缓解定位漂移、模拟行人运动复杂多变的运动模式等问题提供了比较有前景的解决方案. 但是, 它们在现实场景中应用和部署时, 仍然存在一些亟待解决的问题, 例如, 精度、预测效率、设备方向估计、数据稀缺以及场景泛化等. 在本节中, 我们首先介绍当前行人惯性定位技术面临的挑战; 然后, 针对这些问题和挑战提出一些未来的研究趋势.

4.1 面临的挑战

4.1.1 方向估计

设备的方向估计问题是基于神经网络方法的一个主要关注点, 现有的比较主流的研究工作^[26-27] 在将 IMU 测量数据从本体坐标系转换到导航坐标系时, 大多依赖于智能手机中的设备估计方向, 例如游戏旋转向量. 事实上, 被误差源破坏的方向甚至错误的方向都会导致位置积分过程中误差的进一步累积.

4.1.2 异质性问题

由于缺乏可解释性, 神经网络已被贴上“黑盒”的标签. 因此, 它需要更多样性的被试样本和设备样本才能有更好的泛化能力. 但是, 不同的行人被试和设备往往存在异质性: 1) 每一个行人都是不同的个体, 各种特征 (包括身高、腿长、年龄、

性别、体重以及个体的步行习惯) 均会导致步长在不同行人之间存在差异, 例如, 不同身高的行人步长存在差异, 相同身高的不同年龄的行人步长也存在差异. 2) 对于设备而言, 不同供应商提供的不同类型传感器的设备也存在很大的异质性, 例如, 行人惯性定位大多采用智能手机来进行数据采集, 不同类型的智能手机内置的传感器也是不同类型的, 其对应的传感器校准也存在差异, 不一致的惯性数据导致定位误差出现波动, 对模型的泛化能力和自适应能力也提出更高的要求.

4.1.3 场景复杂性

现有的大多数神经惯性定位算法是在室内、建筑内进行实验的, 运动路径和简单的场景使得模型更容易学习到可以回归状态的惯性特征表示. 但是, 在实际环境中, 复杂的障碍物摆放、移动的行人或者车辆会让定位精度急剧下降. 另外, 现有的数据集都会限定行人的运动模式和设备的携带方式, 采集数据的过程中行人也不需要快速、频繁地改变步行航向和模式. 事实上, 人体的灵活结构进一步增加了对惯性传感器捕获的运动信号的复杂度. 多样化的行走模式以及这些传感器捕获的不同携带方式的运动信号包含了来自手、包和口袋的混合信号, 使得准确推理行人状态变得困难.

4.1.4 真值不“真”性

基于神经网络的方法表现卓越在很大程度上依赖于大量的高质量有监督数据. 然而, 获取此类数据可能既昂贵又耗时, 同时真实标签的标注也面临一定挑战: 1) 现有的数据集如 RoNIN、RIDI 等大多基于视觉惯性里程计来采集, 该算法本身也具备漂移的现象, 即使可以使用回环检测矫正误差, 但是也会存在时漂时准的现象, 这会导致真值不“真”; 2) 噪声源的影响也会使得数据在标注过程中出错; 3) 在紧急情况或不寻常的活动模式 (例如横向行走或向后行走) 获取注释数据是具有挑战性的, 这会导致另一个挑战, 称为类别不平衡. 当使用不平衡的数据集时, 模型预测结果偏向于训练样本较多的类别, 忽略样本较少的类别. 对于行人惯性定位, 解决真值不真和类不平衡的问题迫在眉睫.

4.1.5 效率与计算资源约束

让神经网络更深更大是提高模型性能的关键手段, 但是海量的参数不可避免地会增加计算开销和推理时间, 同时需要 GPU 进行并行训练和测试. 由于智能手机/手环等设备的核心功能是通信, 位置感知只是一个额外的服务, 所以它们不能使用全部计算资源和存储空间进行状态预测. 因此, 基于深度神经网络的惯性定位模型的最优结构需要考

虑高性能、轻量化、可部署性等因素。

4.2 未来的研究趋势

在未来,基于神经网络的行人惯性定位模型应该具有自学习和适应能力,并能提供高效率、高精度的定位服务。本节概述了基于神经网络的行人惯性定位几个可能的未来研究趋势,如下所示:

4.2.1 数据集基准

现有的深度学习技术通常基于大量的数据进行训练和测试。因此,数据集需要符合几个基本的要求: 1) 包含各种各样的行人运动模式且模式种类相对平衡; 2) 包含各种场景,如室内室外以及各种复杂环境因素; 3) 包含以各种各样携带方式的移动设备采集的数据,例如多种品牌的智能手机、智能手环、AR/VR 设备等。未来,一个广泛采用的涵盖了各种应用场景的数据集基准,将有利于基于神经网络的行人惯性定位技术的发展。

4.2.2 自学习与自适应方法

由于行人运动的高度复杂性,每一时刻的运动模式几乎都是独一无二的,因此模型的泛化能力和自适应能力变得尤为重要。另外,此方向的多领域数据集会存在训练数据和测试数据的领域差异导致模型定位精度的下降。据此,可以采用以下几种思路来解决此问题: 1) 利用数据增强^[86-87]来增加训练数据的多样性和数量,从而提升模型的性能和泛化能力。它通常通过对原始数据进行一系列的变换或扩展来生成新的训练样本。2) 可以采用增量学习^[88]和在线学习^[89]在定位过程中自适应地估计和补偿传感器误差。3) 目前,迁移学习^[90]也是提高模型泛化能力的常用手段,可以将模型初始化为从源任务学到的知识,并在惯性定位任务上进行微调。4) 通过利用物理/几何规则或其他传感器(如相机)的信息,基于学习的惯性定位模型进行自监督或者无监督训练^[91],使移动设备能够终生从数据中学习。

4.2.3 轻量化网络设计

算法研究的最终目的是为了将算法部署在移动终端上并为人类的日常生活提供便利精准的定位服务,然而移动终端设备不能使用全部计算资源和存储空间进行状态预测。因此,我们需要一个更加轻量级、高效和精准的模型来实现移动终端的实时预测。为了实现这一目标,可以引入一些模型压缩技术,如知识蒸馏^[92],通过训练一个大型的、高性能的模型(通常称为教师模型),然后将其所学到的知识迁移给一个更小、更轻量级的模型(通常称为学生模型)。以此减少模型的大小和计算复杂度,同时保

持模型的性能或提升其泛化能力。另外,模型剪枝技术^[93]也可以通过去除神经网络中不必要的连接或节点来减少参数数量和模型大小。

4.2.4 基于深度学习的多源融合定位

目前,行人惯性定位很难长时间长距离地提供理想的定位精度,如何有效地提高预测精度依旧是研究人员的关注热点。移动设备内置的传感器往往不止一种,例如,智能手机包含 GPS、陀螺仪、加速度计、磁力计、5G、WiFi 和相机等传感器。因此,基于深度学习的融合定位方法也可能是未来提高行人定位的主流方法之一。其中,多传感器信号之间不同模态的对齐^[94-96]、时间同步^[97]、模态融合以及融合策略^[98-102]都是多源融合定位需要攻克的难题。

References

- Gao R P, Xiao X, Zhu S L, Xing W W, Li C, Liu L, et al. Glow in the dark: Smartphone inertial odometry for vehicle tracking in GPS blocked environments. *IEEE Internet of Things Journal*, 2021, 8(16): 12955-12967
- Herrera E, Kaufmann H, Secue J, Quirós R, Fabregat G. Improving data fusion in personal positioning systems for outdoor environments. *Information Fusion*, 2013, 14(1): 45-56
- Roy P, Chowdhury C. A survey on ubiquitous WiFi-based indoor localization system for smartphone users from implementation perspectives. *CCF Transactions on Pervasive Computing and Interaction*, 2022, 4(3): 298-318
- Wang R R, Li Z H, Luo H Y, Zhao F, Shao W H, Wang Q. A robust Wi-Fi fingerprint positioning algorithm using stacked denoising autoencoder and multi-layer perceptron. *Remote Sensing*, 2019, 11(11): Article No. 1293
- Mur-Artal R, Montiel J, Tardós J. ORB-SLAM: A versatile and accurate monocular SLAM system. *IEEE Transactions on Robotics*, 2015, 31(5): 1147-1163
- Szyc K, Nikodem M, Zdunek M. Bluetooth low energy indoor localization for large industrial areas and limited infrastructure. *Ad Hoc Networks*, 2023, 139: Article No. 103024
- Yu N, Zhan X H, Zhao S N, Wu Y F, Feng R J. A precise dead reckoning algorithm based on Bluetooth and multiple sensors. *IEEE Internet of Things Journal*, 2018, 5(1): 336-351
- Chen C H, Zhao P J, Lu C X, Wang W, Markham A, Trigoni N. Deep-learning-based pedestrian inertial navigation: Methods, data set, and on-device inference. *IEEE Internet of Things Journal*, 2020, 7(5): 4431-4441
- Gowda M, Dhekne A, Shen S, Choudhury R, Yang X, Yang L, et al. Bringing IoT to sports analytics. In: Proceedings of the 14th USENIX Conference on Networked Systems Design and Implementation. Boston, USA: USENIX, 2017. 499-513
- Pan Xian-Fei, Mu Hua, Hu Xiao-Ping. A survey of autonomous navigation technology for individual soldier. *Navigation Positioning and Timing*, 2018, 5(1): 1-11 (潘献飞, 穆华, 胡小平. 单兵自主导航技术发展综述. 导航定位与授时, 2018, 5(1): 1-11)
- Guo Xiao-Kuan, Yue Pi-Yu, An Wei-Lian. Distance-inertial guidance of the launch vehicle. *Acta Automatica Sinica*, 1984, 10(4): 361-364 (郭孝宽, 岳丕玉, 安维廉. 运载火箭的距离惯性制导. 自动化学报, 1984, 10(4): 361-364)
- Zhou B D, Wu P, Zhang X, Zhang D J, Li Q Q. Activity semantics-based indoor localization using smartphones. *IEEE Sensors Journal*, 2024, 24(7): 11069-11079
- Barshan B, Durrant-Whyte H F. Inertial navigation systems

- for mobile robots. *IEEE Transactions on Robotics and Automation*, 1995, **11**(3): 328–342
- 14 Wang Wei. Status and development trend of inertial technology. *Acta Automatica Sinica*, 2013, **39**(6): 723–729 (王巍. 惯性技术研究现状及发展趋势. *自动化学报*, 2013, **39**(6): 723–729)
- 15 Dong Ming-Tao, Cheng Jian-Hua, Zhao Lin, Liu Ping. Perspectives on performance evaluation method for inertial integrated navigation system. *Acta Automatica Sinica*, 2022, **48**(10): 2361–2373 (董铭涛, 程建华, 赵琳, 刘萍. 惯性组合导航系统性能评估方法研究进展. *自动化学报*, 2022, **48**(10): 2361–2373)
- 16 Xu R. Research and Application on Navigation Algorithm of Pedestrian Navigation System [Master thesis], Nanjing University of Aeronautics and Astronautics, China, 2008. (许睿. 行人导航系统算法研究与应用实现 [硕士学位论文], 南京航空航天大学, 中国, 2008.)
- 17 Puyol M, Bobkov D, Robertson P, Jost T. Pedestrian simultaneous localization and mapping in multistory buildings using inertial sensors. *IEEE Transactions on Intelligent Transportation Systems*, 2014, **15**(4): 1714–1727
- 18 El-Sheimy N, Hou H Y, Niu X J. Analysis and modeling of inertial sensors using Allan variance. *IEEE Transactions on Instrumentation and Measurement*, 2008, **57**(1): 140–149
- 19 Zhuo W P, Li S J, He T L, Liu M Y, Chan S, Ha S, et al. On-line path description learning based on IMU signals from IoT devices. *IEEE Transactions on Mobile Computing*, DOI: [10.1109/TMC.2024.3406436](https://doi.org/10.1109/TMC.2024.3406436)
- 20 Otter D, Medina J, Kalita J. A survey of the usages of deep learning for natural language processing. *IEEE Transactions on Neural Networks and Learning Systems*, 2021, **32**(2): 604–624
- 21 LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*, 2015, **521**(7553): 436–444
- 22 Ru X, Gu N, Shang H, Zhang H. MEMS inertial sensor calibration technology: Current status and future trends. *Micromachines*, 2022, **13**(6): Article No. 879
- 23 Yang H. Research of High-accuracy Pedestrian Navigation Algorithm Based on MEMS Sensors [Master thesis], Xiamen University, China, 2014. (杨辉. 基于 MEMS 传感器的高精度行人导航算法研究 [硕士学位论文], 厦门大学, 中国, 2014.)
- 24 Savage P. Strapdown inertial navigation integration algorithm design part 1: Attitude algorithms. *Journal of Guidance, Control, and Dynamics*, 1998, **21**(1): 19–28
- 25 Savage P. Strapdown inertial navigation integration algorithm design part 2: Velocity and position algorithms. *Journal of Guidance, Control, and dynamics*, 1998, **21**(2): 208–221
- 26 Rao B, Kazemi E, Ding Y, Shila D, Tucker F, Wang L. CTIN: Robust contextual transformer network for inertial navigation. In: Proceedings of the 36th AAAI Conference on Artificial Intelligence. Palo Alto, CA: AAAI, 2022. 5413–5421
- 27 Herath S, Yan H, Furukawa Y. RoNIN: Robust neural inertial navigation in the wild: Benchmark, evaluations, & new methods. In: Proceedings of the IEEE International Conference on Robotics and Automation (ICRA). Paris, France: IEEE, 2020. 3146–3152
- 28 Wang B, Liu X, Yu B, Jia R, Gan X. Pedestrian dead reckoning based on motion mode recognition using a smartphone. *Sensors*, 2018, **18**(6): Article No. 1811
- 29 Li W, Wang Y, Shao Y, Hu G, Li D. TrackPuzzle: Efficient registration of unlabeled PDR trajectories for learning indoor route graph. *Future Generation Computer Systems*, 2023, **149**: 171–183
- 30 Skog I, Handel P, Nilsson J, Rantakokko J. Zero-velocity detection—An algorithm evaluation. *IEEE Transactions on Bio-medical Engineering*, 2010, **57**(11): 2657–2666
- 31 Zhang Lun-Dong, Lu Xiao-Hui, Li Jun-Zheng, He Mai-Hang. The key technologies and development of pedestrian navigation based on ZUPT. *Navigation Positioning and Timing*, 2020, **7**(3): 141–149 (张伦东, 卢晓慧, 李军正, 何励航. 基于零速修正的行人导航关键技术及研究进展. *导航定位与授时*, 2020, **7**(3): 141–149)
- 32 Harle R. A survey of indoor inertial positioning systems for pedestrians. *IEEE Communications Surveys & Tutorials*, 2013, **15**(3): 1281–1293
- 33 Qian J C, Ma J B, Ying R D, Liu P L, Pei L. An improved indoor localization method using smartphone inertial sensors. In: Proceedings of the International Conference on Indoor Positioning and Indoor Navigation. Montbeliard, France: IEEE, 2013. 1–7
- 34 Ao B K, Wang Y C, Liu H N, Li D Y, Song L, Li J Q. Context impacts in accelerometer-based walk detection and step counting. *Sensors*, 2018, **18**(11): Article No. 3604
- 35 Kang X M, Huang B Q, Qi G D. A novel walking detection and step counting algorithm using unconstrained smartphones. *Sensors*, 2018, **18**(1): Article No. 297
- 36 Weinberg H. Using the ADXL202 in pedometer and personal navigation applications. *Analog Devices AN-602 Application Note*, 2002, **2**(2): 1–6
- 37 Nilsson J, Skog I, Händel P, Hari K. Foot-mounted INS for everybody—an open-source embedded implementation. In: Proceedings of the IEEE/ION Position, Location and Navigation Symposium. Myrtle Beach, USA: IEEE, 2012. 140–145
- 38 Fang L, Antsaklis P J, Montestrucque L A, McMickell M B, Lemmon M, Sun Y S, et al. Design of a wireless assisted pedestrian dead reckoning system—The NavMote experience. *IEEE Transactions on Instrumentation and Measurement*, 2005, **54**(6): 2342–2358
- 39 Goyal P, Ribeiro V J, Saran H, Kumar A. Strap-down pedestrian dead-reckoning system. In: Proceedings of the International Conference on Indoor Positioning and Indoor Navigation. Guimaraes, Portugal: IEEE, 2011. 1–7
- 40 Huang B Q, Qi G D, Yang X K, Zhao L, Zou H. Exploiting cyclic features of walking for pedestrian dead reckoning with unconstrained smartphones. In: Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing. Heidelberg, Germany: ACM, 2016. 374–385
- 41 Wahlström J, Skog I. Fifteen years of progress at zero velocity: A review. *IEEE Sensors Journal*, 2021, **21**(2): 1139–1151
- 42 Chen C H, Pan X F. Deep learning for inertial positioning: A survey. *IEEE Transactions on Intelligent Transportation Systems*, 2024, **25**(9): 10506–10523
- 43 Chen C H, Lu X X, Markham A, Trigoni N. IONet: Learning to cure the curse of drift in inertial odometry. In: Proceedings of the 32nd AAAI Conference on Artificial Intelligence. New Orleans, USA: AAAI, 2018. 6468–6476
- 44 Yu Y, Si X S, Hu C H, Zhang J X. A review of recurrent neural networks: LSTM cells and network architectures. *Neural Computation*, 2019, **31**(7): 1235–1270
- 45 Yao S C, Zhao Y R, Shao H J, Liu S Z, Liu D X, Su L, et al. FastDeepIoT: Towards understanding and optimizing neural network execution time on mobile and embedded devices. In: Proceedings of the 16th ACM Conference on Embedded Networked Sensor Systems. Shenzhen, China: ACM, 2018. 278–291
- 46 Oord A, Dieleman S, Zen H, Simonyan K, Vinyals O, Graves A, et al. Wavenet: A generative model for raw audio. arXiv preprint arXiv: 1609.03499, 2016.
- 47 Jayanth R K, Xu Y S, Wang Z Y, Chatzipantazis E, Gehrig D, Danilidis K. EqNIO: Subequivariant neural inertial odometry. arXiv preprint arXiv: 2408.06321, 2024.
- 48 Chen C H, Miao Y S, Lu C X, Xie L H, Blunsom P, Markham A, et al. MotionTransformer: Transferring neural inertial tracking between domains. In: Proceedings of the 33rd AAAI Conference on Artificial Intelligence. Honolulu, USA: AAAI, 2019. 8009–8016
- 49 Yu L T, Zhang W N, Wang J, Yu Y. SeqGAN: Sequence gen-

- erative adversarial nets with policy gradient. In: Proceedings of the 31st AAAI Conference on Artificial Intelligence. San Francisco, USA: AAAI, 2017. 2852–2858
- 50 Tzeng E, Hoffman J, Saenko K, Darrell T. Adversarial discriminative domain adaptation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA: IEEE, 2017. 7167–7176
- 51 Liu W X, Caruso D, Ilg E, Dong J, Mourikis A I, Daniilidis K, et al. TLIO: Tight learned inertial odometry. *IEEE Robotics and Automation Letters*, 2020, **5**(4): 5653–5660
- 52 Li M Y, Mourikis A I. High-precision, consistent EKF-based visual-inertial odometry. *The International Journal of Robotics Research*, 2013, **32**(6): 690–711
- 53 Wang Y, Cheng H, Wang C, Meng M. Pose-invariant inertial odometry for pedestrian localization. *IEEE Transactions on Instrumentation and Measurement*, 2021, **70**: Article No. 8503512
- 54 Zeinali B, Zanddizari H, Chang M J. IMUNet: Efficient regression architecture for inertial IMU navigation and positioning. *IEEE Transactions on Instrumentation and Measurement*, 2024, **73**: Article No. 2516213
- 55 Cao X Y, Zhou C F, Zeng D D, Wang Y L. RIO: Rotation-equivariance supervised learning of robust inertial odometry. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans, USA: IEEE, 2022. 6614–6623
- 56 Lai R C, Tian Y, Tian J D, Wang J, Li N, Jiang Y. ResMixer: A lightweight residual mixer deep inertial odometry for indoor positioning. *IEEE Sensors Journal*, DOI: [10.1109/JSEN.2024.3443311](https://doi.org/10.1109/JSEN.2024.3443311)
- 57 Wang Y, Cheng H, Meng M. Inertial odometry using hybrid neural network with temporal attention for pedestrian localization. *IEEE Transactions on Instrumentation and Measurement*, 2022, **71**: Article No. 7503610
- 58 Zhu Y Q, Zhang J L, Zhu Y P, Zhang B, Ma W Z. RBCN-Net: A data-driven inertial navigation algorithm for pedestrians. *Applied Sciences*, 2023, **13**(5): Article No. 2969
- 59 Chen B X, Zhang R F, Wang S C, Zhang L Q, Liu Y. Deep-learning-based inertial odometry for pedestrian tracking using attention mechanism and Res2Net module. *IEEE Sensors Letters*, 2022, **6**(11): Article No. 6003804
- 60 Gao S H, Cheng M M, Zhao K, Zhang X Y, Yang M H, Torr P. Res2Net: A new multi-scale backbone architecture. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021, **43**(2): 652–662
- 61 Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez A, et al. Attention is all you need. In: Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach, USA: Curran Associates Inc., 2017. 5998–6008
- 62 Brothie J, Li W C, Greentree A D, Kealy A. RIOT: Recursive inertial odometry transformer for localisation from low-cost IMU measurements. *Sensors*, 2023, **23**(6): Article No. 3217
- 63 Herath S, Caruso D, Liu C, Chen Y F, Furukawa Y. Neural inertial localization. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans, USA: IEEE, 2022. 6604–6613
- 64 Sun S, Melamed D, Kitani K. IDOL: Inertial deep orientation-estimation and localization. In: Proceedings of the 35th AAAI Conference on Artificial Intelligence. Palo Alto, CA: AAAI, 2021. 6128–6137
- 65 Wang Y, Cheng H, Meng M. Pedestrian motion tracking by using inertial sensors on the smartphone. In: Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Las Vegas, USA: IEEE, 2020. 4426–4431
- 66 Shao W H, Luo H Y, Zhao F, Wang C, Crivello A, Tunio M. DePedo: Anti periodic negative-step movement pedometer with deep convolutional neural networks. In: Proceedings of the IEEE International Conference on Communications (ICC). Kansas City, USA: IEEE, 2018. 1–6
- 67 Ren P, Elyasi F, Manduchi R. Smartphone-based inertial odometry for blind walkers. *Sensors*, 2021, **21**(12): Article No. 4033
- 68 Han K, Yu S M, Ko S W, Kim S L. Waveform-guide transformation of IMU measurements for smartphone-based localization. *IEEE Sensors Journal*, 2023, **23**(17): 20379–20389
- 69 Gu F Q, Khoshelham K, Yu C Y, Shang J G. Accurate step length estimation for pedestrian dead reckoning localization using stacked autoencoders. *IEEE Transactions on Instrumentation and Measurement*, 2019, **68**(8): 2705–2713
- 70 Gehring J, Miao Y J, Metze F, Waibel A. Extracting deep bottleneck features using stacked auto-encoders. In: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing. Vancouver, Canada: IEEE, 2013. 3377–3381
- 71 Klein I, Asraf O. StepNet—Deep learning approaches for step length estimation. *IEEE Access*, 2020, **8**: 85706–85713
- 72 Wang Q, Ye L L, Luo H Y, Men A D, Zhao F, Huang Y. Pedestrian stride-length estimation based on LSTM and denoising autoencoders. *Sensors*, 2019, **19**(4): Article No. 840
- 73 Bo F, Li J, Wang W B. Mode-independent stride length estimation with IMUs in smartphones. *IEEE Sensors Journal*, 2022, **22**(6): 5824–5833
- 74 Im C, Eom C, Lee H, Jang S, Lee C. Deep LSTM-based multimodal pedestrian dead reckoning system for indoor localization. In: Proceedings of the International Conference on Electronics, Information, and Communication (ICEIC). Jeju, South Korea: IEEE, 2022. 1–2
- 75 Huang Y, Zeng Q H, Lei Q Y, Chen Z J, Sun K C. Smartphone heading correction method based on LSTM neural network. In: Proceedings of China Satellite Navigation Conference. Beijing, China: Springer, 2022. 415–425
- 76 Wang Q, Luo H Y, Ye L L, Men A D, Zhao F, Huang Y, et al. Pedestrian heading estimation based on spatial transformer networks and hierarchical LSTM. *IEEE Access*, 2019, **7**: 162309–162322
- 77 Jaderberg M, Simonyan K, Zisserman A, Kavukcuoglu K. Spatial transformer networks. In: Proceedings of the 28th International Conference on Neural Information Processing Systems. Montreal, Canada: MIT Press, 2015. 2017–2025
- 78 Manos A, Hazan T, Klein I. Walking direction estimation using smartphone sensors: A deep network-based framework. *IEEE Transactions on Instrumentation and Measurement*, 2022, **71**: Article No. 2501112
- 79 Asraf O, Shama F, Klein I. PDRNet: A deep-learning pedestrian dead reckoning framework. *IEEE Sensors Journal*, 2022, **22**(6): 4932–4939
- 80 Wagstaff B, Kelly J. LSTM-based zero-velocity detection for robust inertial navigation. In: Proceedings of the International Conference on Indoor Positioning and Indoor Navigation (IPIN). Nantes, France: IEEE, 2018. 1–8
- 81 Yu X G, Liu B, Lan X Y, Xiao Z L, Lin S S, Yan B, et al. AZUPT: Adaptive zero velocity update based on neural networks for pedestrian tracking. In: Proceedings of the IEEE Global Communications Conference (GLOBECOM). Waikoloa, USA: IEEE, 2019. 1–6
- 82 Yan H, Shan Q, Furukawa Y. RIDI: Robust IMU double integration. In: Proceedings of the 15th European Conference on Computer Vision (ECCV). Munich, Germany: Springer, 2018. 621–636
- 83 Schubert D, Goll T, Demmel N, Usenko V, Stückler J, Cremers D. The TUM VI benchmark for evaluating visual-inertial odometry. In: Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Madrid, Spain: IEEE, 2018. 1680–1687
- 84 Chen C H, Zhao P J, Lu C X, Wang W, Markham A, Trigoni N. OxIOD: The dataset for deep inertial odometry. arXiv preprint arXiv: 1809.07491, 2018.

- 85 Liu F, Ge H Y, Tao D, Gao R P, Zhang Z. Smartphone-based pedestrian inertial tracking: Dataset, model, and deployment. *IEEE Transactions on Instrumentation and Measurement*, 2024, **73**: Article No. 2504713
- 86 Shorten C, Khoshgoftaar T M. A survey on image data augmentation for deep learning. *Journal of Big Data*, 2019, **6**(1): Article No. 60
- 87 Shorten C, Khoshgoftaar T M, Furht B. Text data augmentation for deep learning. *Journal of Big Data*, 2021, **8**(1): Article No. 101
- 88 Wu Y, Chen Y P, Wang L J, Ye Y C, Liu Z C, Guo Y D, et al. Large scale incremental learning. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, USA: IEEE, 2019. 374–382
- 89 Zou H, Lu X X, Jiang H, Xie L H. A fast and precise indoor localization algorithm based on an online sequential extreme learning machine. *Sensors*, 2015, **15**(1): 1804–1824
- 90 Iman M, Arabnia H R, Rasheed K. A review of deep transfer learning and recent advancements. *Technologies*, 2023, **11**(2): Article No. 40
- 91 Chen Kang-Xin, Zhao Jie-Yu, Chen Hao. A vector spherical convolutional network based on self-supervised learning. *Acta Automatica Sinica*, 2023, **49**(6): 1354–1368 (陈康鑫, 赵杰煜, 陈豪. 一种基于自监督学习的矢量球面卷积网络. *自动化学报*, 2023, **49**(6): 1354–1368)
- 92 Gou J P, Yu B S, Maybank S J, Tao D C. Knowledge distillation: A survey. *International Journal of Computer Vision*, 2021, **129**(6): 1789–1819
- 93 Blalock D W, Gonzalez J J, Frankle J, Gutttag J. What is the state of neural network pruning? In: Proceedings of the 3rd Machine Learning and Systems. Austin, USA: mlsys.org, 2020. 129–146
- 94 Zhao W Y, Zhou D, Cao B Q, Zhang K, Chen J J. Adversarial modality alignment network for cross-modal molecule retrieval. *IEEE Transactions on Artificial Intelligence*, 2024, **5**(1): 278–289
- 95 Freydin M, Segol N, Sfaradi N, Eweida A, Or B. Deep learning for inertial sensor alignment. *IEEE Sensors Journal*, 2024, **24**(10): 17282–17290
- 96 Aslan M F, Durdu A, Sabanci K. Visual-Inertial Image-Odometry Network (VIIONet): A Gaussian process regression-based deep architecture proposal for UAV pose estimation. *Measurement*, 2022, **194**: Article No. 111030
- 97 Nilsson J O, Händel P. Time synchronization and temporal ordering of asynchronous sensor measurements of a multi-sensor navigation system. In: Proceedings of the IEEE/ION Position, Location and Navigation Symposium. Indian Wells, USA: IEEE, 2010. 897–902
- 98 Chen C H, Rosa S, Miao Y S, Lu C X, Wu W, Markham A, et al. Selective sensor fusion for neural visual-inertial odometry. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, USA: IEEE, 2019. 10542–10551
- 99 Li Shuai-Xin, Li Guang-Yun, Wang Li, Yang Xiao-Tian. LiDAR/IMU tightly coupled real-time localization method. *Acta Automatica Sinica*, 2021, **47**(6): 1377–1389 (李帅鑫, 李广云, 王力, 杨啸天. LiDAR/IMU 紧耦合的实时定位方法. *自动化学报*, 2021, **47**(6): 1377–1389)
- 100 Almalioglu Y, Turan M, Saputra M, Gusmão P, Markham A, Trigoni N. SelfVIO: Self-supervised deep monocular visual-inertial odometry and depth estimation. *Neural Networks*, 2022, **150**: 119–136
- 101 Zhou P, Wang H, Gravina R, Sun F M. WIO-EKF: Extended Kalman filtering-based Wi-Fi and inertial odometry fusion method for indoor localization. *IEEE Internet of Things Journal*, 2024, **11**(13): 23592–23603
- 102 Li J Y, Pan X K, Huang G, Zhang Z Y, Wang N, Bao H J, et al. RD-VIO: Robust visual-inertial odometry for mobile aug-

mented reality in dynamic environments. *IEEE Transactions on Visualization and Computer Graphics*, 2024, **30**(10): 6941–6955



李岩 中山大学系统科学与工程学院博士研究生. 2022 年获得昆明理工大学硕士学位. 主要研究方向为神经惯性定位和姿态估计.

E-mail: liyan377@mail2.sysu.edu.cn

(**LI Yan** Ph.D. candidate at the School of Systems Science and Engineering,

Sun Yat-Sen University. He received his master degree from Kunming University of Science and Technology in 2022. His research interest covers neural inertial positioning and pose estimation.)



施忠臣 军事科学院国防科技创新研究院助理研究员. 2021 年获得国防科技大学博士学位. 主要研究方向为机器人视觉和姿态估计.

E-mail: shizhongchen@buaa.edu.cn

(**SHI Zhong-Chen** Assistant researcher at the Defense Innovation

Institute, Academy of Military Sciences. He received his Ph.D. degree from National University of Defense Technology in 2021. His research interest covers robot vision and pose estimation.)



侯燕青 中山大学系统科学与工程学院副教授. 2016 年获得国防科技大学博士学位. 主要研究方向为卫星导航定位和多源融合导航.

E-mail: hoyq9@mail.sysu.edu.cn

(**HOU Yan-Qing** Associate professor at the School of Systems Science

and Engineering, Sun Yat-Sen University. He received his Ph.D. degree from National University of Defense Technology in 2016. His research interest covers satellite navigation positioning and multi-source fusion navigation.)



戚煜华 中山大学系统科学与工程学院副研究员. 2020 年获得北京理工大学博士学位. 主要研究方向为同时定位与建图.

E-mail: qiyh8@mail.sysu.edu.cn

(**QI Yu-Hua** Associate researcher at the School of Systems Science

and Engineering, Sun Yat-Sen University. He received his Ph.D. degree from Beijing Institute of Technology in 2020. His main research interest is simultaneous positioning and mapping.)



谢良 军事科学院国防科技创新研究院副研究员. 2018 年获得国防科技大学博士学位. 主要研究方向为计算机视觉, 人机交互和混合现实.

E-mail: xielnudt@gmail.com

(**XIE Liang** Associate researcher at the Defense Innovation Institute, Academy of Military Sciences. He received his Ph.D. degree from National University of Defense Technology in 2018. His research interest covers computer vision, human-machine interaction and mixed reality.)



陈伟 军事科学院国防科技创新研究院助理研究员. 2022 年获得伯明翰大学博士学位. 主要研究方向为姿态估计.

E-mail: wei.chen.ai@outlook.com

(**CHEN Wei** Assistant researcher at the Defense Innovation Institute, Academy of Military Sciences. He received his Ph.D. degree from the University of Birmingham in 2022. His main research interest is pose estimation.)



陈洪波 中山大学系统科学与工程学院教授. 2007 年获得哈尔滨工业大学博士学位. 主要研究方向为复杂系统的建模、仿真和分析, 智能无人系统以及航空航天飞行器的设计.

E-mail: chenhongbo@mail.sysu.edu.cn

(**CHEN Hong-Bo** Professor at the School of Systems Science and Engineering, Sun Yat-

Sen University. He received his Ph.D. degree from the Harbin Institute of Technology in 2007. His research interest covers modeling, simulation and analysis of complex systems, intelligent unmanned systems and design of aerospace vehicles.)



闫野 军事科学院国防科技创新研究院研究员. 2000 年获得国防科技大学博士学位. 主要研究方向为人机交互和混合现实. 本文通信作者.

E-mail: yanye1971@sohu.com

(**YAN Ye** Researcher at the Defense Innovation Institute, Academy of Military Sciences. He received his Ph.D. degree from National University of Defense Technology in 2000. His research interest covers human-machine interaction and mixed reality. Corresponding author of this paper.)



印二威 军事科学院国防科技创新研究院研究员. 2015 年获得国防科技大学博士学位. 主要研究方向为脑机接口和智能人机交互技术.

E-mail: yinerwei1985@gmail.com

(**YIN Er-Wei** Researcher at the Defense Innovation Institute, Academy of Military Sciences. He received his Ph.D. degree from National University of Defense Technology in 2015. His research interest covers brain-computer interfaces and intelligent human-machine interaction technologies.)