

基于双尺度门控 Mamba-CNN 可微卡尔曼滤波器融合定位算法

蒋锐, 张腾, 张天驰, 沈子航, 王小明, 徐友云

(南京邮电大学通信与信息工程学院, 江苏 南京 210003)

摘要: 针对现有可微卡尔曼滤波 (DKF) 方法依赖循环神经网络而导致长程记忆能力受限、难以同时抑制传感器长时零偏漂移与量测瞬态冲击, 以及在复杂城市环境中面临 GPS 长时间拒止时状态估计易发散的问题, 本文提出一种双尺度门控 Mamba-CNN 可微卡尔曼滤波器 (Dual-Scale Gated Mamba-CNN Differentiable Kalman Filter, DSG-MC DKF) 融合定位算法。该方法将 IMU 序列、观测新息与断联掩码输入双尺度 Mamba-CNN 网络, 并通过数据驱动增益预测调节状态的更新过程, 从而降低对外部逐时刻姿态旋转解算结果和精确人工运动学模型的直接依赖; 同时引入显式断联掩码门控机制, 并结合 Mamba 与 1D-CNN 双尺度并行结构, 实现长程趋势特征与短时瞬态特征的自适应融合。实验结果表明, 相较于传统 GRU 架构, 所提模型在提升定位精度的同时将单步推理延迟降低了 80%; 在长达 10 s 的 GPS 完全拒止场景下, 仍能够较好维持载体运动轨迹的物理拓扑保真度。

关键词: 可微卡尔曼滤波; 融合定位; 结构化状态空间模型; GNSS 拒止

中图分类号: TP309

文献标志码: A

doi: 10.11959/j.issn.1000

Fusion Localization Based on a Dual-Scale Gated Mamba-CNN Differentiable Kalman Filter

Jiang Rui, Zhang Teng, Zhang Tianchi, Shen Zihang, Wang Xiaoming, Xu Youyun

College of Telecommunications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China

Abstract: To address the limitations of existing differentiable Kalman filter (DKF) methods, which commonly rely on recurrent neural-network architectures and therefore suffer from limited long-range memory, difficulty in simultaneously suppressing long-term sensor bias drift and transient measurement shocks, and degraded state-estimation stability under prolonged GPS outages in complex urban environments, this paper proposes a dual-scale gated Mamba-CNN differentiable Kalman filter (DSG-MC DKF) for fusion localization. The proposed method removes the explicit control-input term from the prediction equation and feeds IMU sequences, measurement innovations, and outage masks into a dual-scale Mamba-CNN network for data-driven gain prediction, thereby reducing the direct reliance on externally computed frame-rotation results and precisely specified kinematic models. Moreover, an explicit outage-mask gating mechanism is introduced, and a dual-scale parallel architecture consisting of Mamba and 1D-CNN branches is developed to decouple long-term bias compensation from high-frequency noise suppression. Experimental results show that, compared with the conventional GRU-based DKF, the proposed model improves localization robustness and accuracy while reducing single-step inference latency by approximately 80%. Furthermore, during a complete GPS outage lasting up to 10 s, the proposed model can still effectively preserve the kinematic and topological consistency of the vehicle trajectory.

Key words: differentiable Kalman filter, fusion localization, structured state space model, GNSS outage

收稿日期: XXXX-XX-XX; 修回日期: XXXX-XX-XX

通信作者: 蒋锐, j_ray@njupt.edu.cn

基金项目: 国家自然科学基金资助项目 (No.62371246)

Foundation Items: The National Natural Science Foundation of China (No.62371246)

0 引言

状态估计是自动驾驶、无人机及智能移动机器人感知与理解外部环境的核心基石,其估计结果的精度与鲁棒性直接决定了上层路径规划与运动控制系统运行的安全性和高效性。在过去半个世纪中,扩展卡尔曼滤波(Extended Kalman Filter, EKF)及其衍生算法在状态估计领域长期占据主流地位。此类传统模型驱动(Model-driven)方法具备严密的数理逻辑与高度的可解释性,因而在诸如惯性测量单元(IMU)与全球定位系统(GPS)等多传感器融合定位领域长期发挥着主导作用。经典的状态估计理论长期采用模型驱动范式,其核心方法为贝叶斯滤波(Bayesian Filtering)。自卡尔曼滤波(Kalman Filter, KF)提出以来,为了克服非线性系统在状态估计中的计算难题,研究人员相继提出了一系列演进的算法,如扩展卡尔曼滤波(EKF)、无迹卡尔曼滤波(UKF)以及容积卡尔曼滤波(CKF)^[1-3]。上述方法借助泰勒级数展开、无迹变换、容积积分规则等数学工具,以解析或数值近似的方式应对系统的非线性特征,在导航定位等领域取得了里程碑式的成功^[4-8]。

然而,上述传统滤波算法存在两个明显的局限性^{[9][10]}。其一,此类算法本质上通常建立在严格的线性化近似与高斯白噪声假设之上,但在现实生活中,这些理想化条件往往难以满足。具体而言,其状态推演过程高度依赖于所用的数学模型,当系统呈现高阶非线性动态特性或发生严重的动力学模型失配时,极易导致状态估计产生显著偏差。而且此类误差一旦随滤波迭代传播,滤波器将面临着发散的风险。其二,在面临城市峡谷、地下隧道或密集林区等GPS信号完全消失的典型场景时,由于传统方法缺乏针对复杂环境的自适应调节机制,在失去外部绝对量测信息的条件下,系统难以有效抑制传感器的发散漂移,最终导致定位结果迅速劣化甚至彻底失效。

为突破上述瓶颈,部分学者尝试引入纯数据驱动(Data-driven)方法,即利用深度神经网络直接对传感器原始时序数据进行端到端的非线性回归^[11-19]。此类方法通常摒弃了传统滤波器中的观测方程与增益计算,转而在数据输入端进行直接拼接,或构建独立分支提取高维表征后再利用全连接层及注意力机制(Attention Mechanism)进行特征

级关联,以此隐式地学习异构传感器间的非线性映射关系。例如,经典的VINet模型通过CNN与LSTM分别提取视觉和惯性特征后进行直接拼接,首次实现了完全不依赖复杂卡尔曼滤波方程与物理标定的端到端位姿估计^[11]。近年来,长序列时间建模方法在多变量时序预测任务中也取得了较快发展,为传感器序列特征提取提供了新的参考。例如,Temporal Fusion Transformer(TFT)^[20]通过变量选择网络、门控结构与可解释注意力机制,实现了多变量时间序列的多步预测;FEDformer^[21]将季节-趋势分解与频域增强机制引入Transformer架构,以捕获长序列中的全局周期结构与趋势变化;TimeMixer^[22]从多尺度分解与混合角度刻画不同时间尺度下的微观波动与宏观趋势;TimeMachine和S-Mamba^{[23][24]}等方法进一步将选择性状态空间模型引入长期时间序列预测任务,以提升长程依赖建模效率。此外,近期针对长期时间序列预测的评测研究^[25]也指出,不同模型在不同数据集、预测长度和实验设置下的性能差异较大,单一复杂模型并不总能在所有场景中保持绝对优势。

上述研究表明,深度网络在复杂时序特征提取、长期依赖建模和多变量相关性挖掘方面具有较强能力。然而,此类方法本质上仍多属于纯数据驱动的直接序列映射框架,通常缺少卡尔曼滤波中“预测—更新”递归闭环所提供的物理约束和状态一致性保证。尤其在城市峡谷、隧道等复杂导航场景中,当GPS长时间拒止、观测信号突然恢复或传感器噪声呈现强非平稳特性时,纯数据驱动方法仍可能出现轨迹拓扑失真、恢复期尖峰或状态估计发散等问题^[17-19],因而难以直接满足安全关键型融合定位系统对稳定性、鲁棒性和可解释性的要求。

为进一步兼顾深度网络的非线性表达能力与滤波模型的物理一致性,部分学者融合上述两种方法的思路,提出了可微卡尔曼滤波(Differentiable Kalman Filter, DKF),一类兼具物理机制与数据驱动优势的方法^{[26][27]}。其核心思想在于保留经典卡尔曼滤波的“预测—更新”递归闭环框架,同时借助深度神经网络的非线性表达能力,根据历史观测、新息及状态演化特征预测用于状态更新的增益矩阵,从而减少对显式协方差递推和矩阵求逆过程的依赖。在该领域,Revach等人提出的Kalman-Net是具有代表性的研究之一。该方法以实时新息

及状态演化特征为输入,利用门控循环单元(GRU)动态预测数据驱动增益矩阵,并将其作用于标准新息更新方程,在多类仿真环境中取得了优于传统扩展卡尔曼滤波(EKF)的状态估计结果^[27]。该类网络输出的增益并非通过显式协方差递推解析获得,因此属于面向特定训练任务的增益近似,并不自动继承经典卡尔曼增益的统计最优性。尽管上述算法在受控实验环境中表现出良好的状态估计性能,但在面对复杂且动态多变的真实场景时,仍暴露出三个缺陷:首先是长序列建模能力匮乏。基于RNN架构的网络在沿时间步反向传播的过程中,极易遭受梯度消失或信息衰减的困扰,导致其长期记忆容量严重受限^{[28][29]}。这一缺陷致使其难以有效捕获并消除传感器中随时间缓慢累积的低频零偏漂移(Bias Drift),无法应对长周期的误差累计;其次是抗瞬态信号冲击能力薄弱。在GPS信号历经长时间的消失后骤然恢复的情况下,其数据往往会产生剧烈的阶跃跳变。这种高频瞬态冲击极易导致RNN内部隐状态(Hidden State)的崩溃,进而引发状态估计结果在信号恢复瞬间出现剧烈波动,即典型的“恢复期尖峰(Recovery Spike)”现象;最后,部分融合定位方法在状态预测阶段需要利用外部姿态解算结果,将载体坐标系下的IMU量测逐时刻旋转至全局坐标系,并进一步通过人工构建的运动学模型形成控制输入。该处理方式能够为状态预测提供明确的物理约束,但其性能会受到姿态估计精度、传感器安装误差及模型参数失配的影响,且增加了外部姿态解算模块与滤波器之间的耦合。基于此,本文不再将外部逐时刻坐标旋转结果作为预测方程中的显式控制输入,而是将IMU时序信息引入数据驱动增益预测网络,以降低外部姿态解算误差对状态预测过程的直接影响。需要说明的是,该设计并不取消传感器标定、时间同步、初始坐标定义以及必要的地理坐标预处理。

综上所述,本文提出了基于双尺度门控Mamba-CNN可微卡尔曼滤波器(Dual-Scale Gated Mamba-CNN Differentiable Kalman Filter, DSG-MC DKF)融合定位算法。该算法突破了传统序列模型依赖滑动窗口机制的性能瓶颈,将兼具线性计算复杂度与全局感受野的结构化状态空间模型(Structured State Space Model, SSM/Mamba)深度

嵌入至基于物理规则的贝叶斯滤波闭环之中。具体而言,为降低状态估计对外部逐时刻姿态旋转解算结果和精确人工运动学模型的直接依赖,本文在滤波预测方程中采用无显式控制输入的标称状态传播策略,将控制输入矩阵置零,并将IMU序列、观测新息及断联标志输入双尺度Mamba-CNN网络。网络并不直接输出载体坐标系到全局坐标系的变换结果,而是利用历史传感器特征预测数据驱动增益矩阵,并通过卡尔曼新息更新通道缓解姿态误差、传感器偏置和运动学模型失配对状态估计产生的综合影响。

此外,针对传感器量测信号长时间消失的极端情况下,本文算法设计了融合时序偏移逻辑的显式断联掩码(Explicit Disconnection Mask)与特征清洗机制。该机制与负责提取高通瞬态特征的一维因果卷积分支(1D Causal Convolution Branch)并行运作,随后通过可学习的非线性门控层(Gating Layer)实现双尺度特征的动态融合。这一架构设计不仅从底层机制上有效阻断了GPS信号消失期间虚假梯度与异常量测噪声对滤波器状态更新的负面干扰,更使系统具备了在“长程物理运动惯性先验”与“短时瞬态高频观测信息”之间进行自适应平滑切换的动态加权能力。

本文的主要贡献可归纳为以下三个方面:

1) 理论建模层面,本文在可微卡尔曼滤波框架中引入无显式控制输入的标称状态传播策略,并将IMU时序特征、观测新息与断联掩码引入数据驱动增益预测过程,从而降低对外部逐时刻坐标旋转结果和精确人工运动学模型的直接依赖,构建了一种保留卡尔曼预测—更新结构的多传感器融合方法。

2) 在网络架构层面,本文设计了并行的双尺度特征提取模块Mamba与一维CNN分支,并通过引入可学习的动态门控层(Dynamic Gating Layer),有效解耦了针对长程零偏漂移的宏观补偿与针对高频量测噪声的微观过滤机制。

3) 在机理验证与真实场景评估层面,本文在洛伦兹混沌系统、高维线性模型失配和缓变漂移场景中评估了双尺度架构对非线性动态及长程误差的建模能力,并在UrbanNav-HK TST和Whampoa两条未参与训练的真实城市路线中开展零样本测试。实验结果表明,在本文训练分布和测试条件下,所

提方法能够改善传统 GRU 增益网络的状态估计性能,并在最长 10 s 的 GPS 完全拒止实验中较好保持轨迹连续性。该结果反映了模型在所测试运动模式、噪声范围和拒止长度内的迁移能力,不代表其对任意真实场景和任意长度的 GNSS 拒止均具有相同性能。此外,所提模型在统一测试环境下表现出较低的推理延迟,为复杂导航环境下的实时状态估计提供了一种可行方案。

1 扩展卡尔曼滤波

在可微卡尔曼滤波相关研究出现之前,扩展卡尔曼滤波(EKF)长期占据非线性状态估计的主导地位^[1-3]。该方法的核心思想是通过一阶泰勒展开对非线性系统进行局部线性化。考虑如下离散时间非线性状态空间模型

$$x_t = f(x_{t-1}, u_t) + w_t, \quad w_t \sim \mathcal{N}(0, Q) \quad (1)$$

$$y_t = h(x_t) + v_t, \quad v_t \sim \mathcal{N}(0, R) \quad (2)$$

其中 x_t 为 t 时刻的系统状态, y_t 为观测值, $f(\cdot)$ 和 $h(\cdot)$ 分别为状态转移函数与观测函数, m 、 n 、 r 分别表示状态维度、观测维度和输入维度, w_t 、 v_t 为互不相关的高斯噪声,协方差矩阵分别 Q 和 R 。EKF 通过递归的“预测-更新”步骤实现状态估计,预测阶段为

$$P_{t|t-1} = \hat{F}_t P_{t-1|t-1} \hat{F}_t^T + Q \quad (3)$$

$$S_{t|t-1} = \hat{H}_t P_{t|t-1} \hat{H}_t^T + R \quad (4)$$

其中 $\hat{F}_t$ 和 $\hat{H}_t$ 分别为 f 和 h 在估计点处的雅可比矩阵。 $P_{t|t-1}$ 和 $P_{t-1|t-1}$ 分别表示预测协方差矩阵和上一时刻更新后的协方差矩阵, $S_{t|t-1}$ 表示新息协方差

矩阵。

更新阶段为

$$K_t = P_{t|t-1} \hat{H}_t^T S_{t|t-1}^{-1} \quad (5)$$

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + K_t (y_t - \hat{y}_{t|t-1}) \quad (6)$$

其中 K_t 为卡尔曼增益矩阵, $y_t - \hat{y}_{t|t-1}$ 为观测新息,因此, $K_t (y_t - \hat{y}_{t|t-1})$ 可与预测状态 $\hat{x}_{t|t-1}$ 进行维度一致的状态更新。

EKF 的完整计算流包含一阶矩(状态)的显式传播与二阶矩(协方差)的复杂递推,后者涉及雅可比矩阵计算与矩阵求逆操作。尽管 EKF 在理论上

具有严密性,但在实际工程应用中存在固有局限。首先,其强依赖噪声先验,卡尔曼增益 K_t 的计算高度依赖于系统噪声协方差 Q 与观测噪声协方差 R ,而实际场景中这些统计特性往往未知或时变,易引发模型失配。其次,存在局部线性化误差,对于强非线性系统,一阶泰勒展开会引入截断误差,严重时可能导致滤波发散。最后,计算不稳定性问题突出,每一步均需计算创新协方差矩阵的逆,其计算复杂度随状态维数呈立方增长,且数值稳定性较差。经典扩展卡尔曼滤波完整的“预测-更新”计算流如图 1 所示。

2 可微卡尔曼滤波

为突破上述瓶颈,研究者提出了可微卡尔曼滤波(Differentiable Kalman Filter, DKF)^{[26][27]}。其核心思想是在保留 EKF 可解释的“预测——更新”闭环结构的前提下,利用神经网络从历史

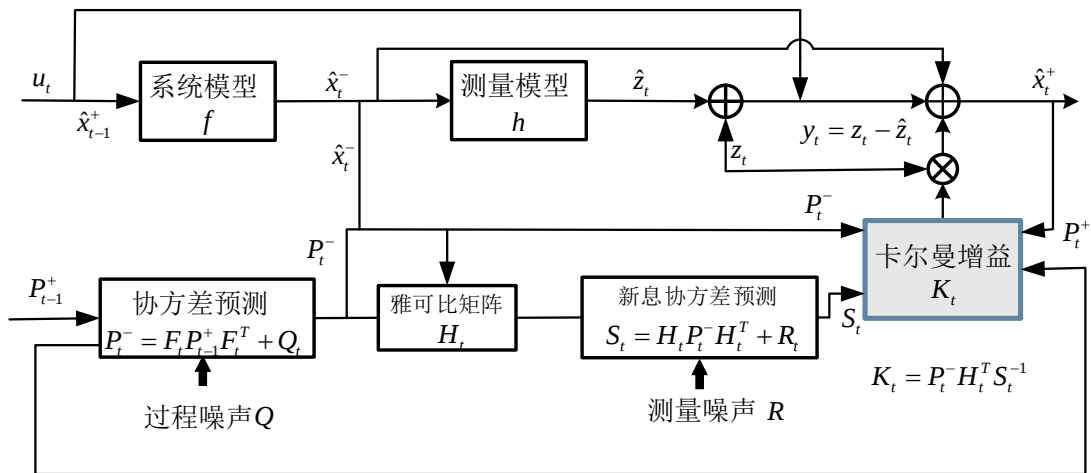


图 1 经典扩展卡尔曼滤波(EKF)计算流程图

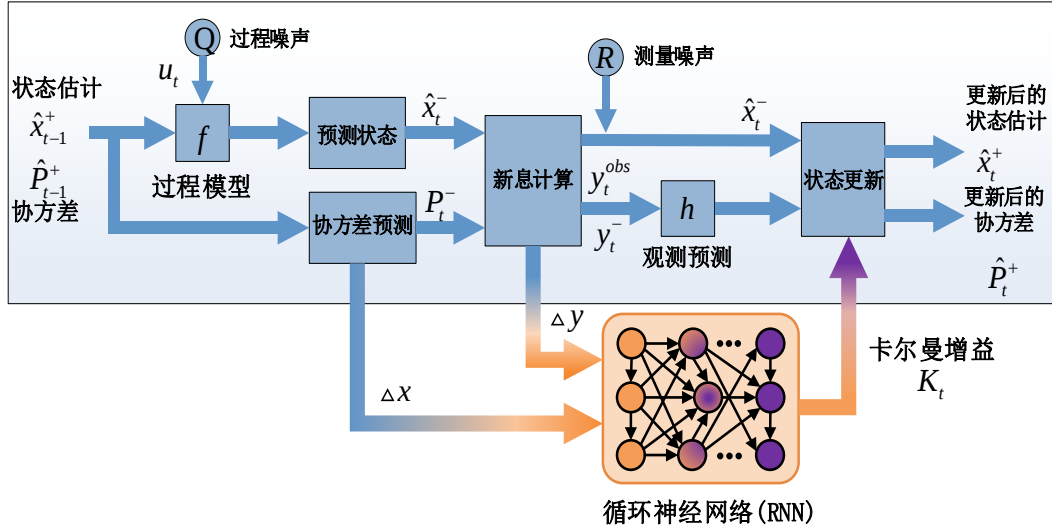


图2 可微卡尔曼滤波架构图

观测、

新息及状态演化特征中学习任务相关的增益矩阵，以减少对显式协方差递推和矩阵求逆过程的依赖。

具体而言,DKF 的预测步骤仍沿用物理模型。

$$\hat{x}_{t|t-1} = f(\hat{x}_{t-1|t-1}), \quad \hat{y}_{t|t-1} = h(\hat{x}_{t|t-1}) \quad (7)$$

但在更新步骤中,卡尔曼增益 K_t 不再由解析协方差计算,而是由一个循环神经网络隐式生成

$$K_t = \mathcal{F}_\theta(\Phi_t, h_{t-1}) \quad (8)$$

其中 $\mathcal{F}_\theta$ 表示由神经网络参数 θ 参数化的卡尔曼增益预测函数, h_{t-1} 为 RNN 的隐状态,用于捕捉时序二阶统计特征; Φ_t 为输入特征,通常构造为与新息和状态演进相关的差分变量。网络输出的卡尔曼增益满足 $K_t \in \mathbb{R}^{m \times n}$,从而保证后续状态更新的维度一致的状态修正量。

$$\Delta y_t = y_t - \hat{y}_{t|t-1} \quad (9)$$

$$\Delta \hat{x}_{t-1} = \hat{x}_{t-1|t-1} - \hat{x}_{t-2|t-2} \quad (10)$$

其中, Δy_t 表示观测新息或观测差分特征, $\Delta \hat{x}_{t-1}$ 表示相邻时刻状态估计的变化量,二者分别用于表征观测偏差和状态演化趋势。

获得 K_t 后,状态更新仍遵循标准形式

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + K_t \Delta y_t \quad (11)$$

其中, $\hat{x}_{t|t-1}$ 为预测状态, $\hat{x}_{t|t}$ 为更新后的状态估计。

该算法利用可学习的 RNN 模块近似历史时序特征到增益矩阵之间的非线性映射,从而减少了对准确噪声先验、显式协方差递推和矩阵求逆过程的

依赖,同时仍保留预测—更新递归结构对状态估计过程的约束,其整体架构如图2所示。

DKF 的另一关键优势在于其端到端可微训练特性。由于卡尔曼更新方程对 K_t 解析可导,可采用真实状态 x_t 与估计状态 $\hat{x}_{t|t}$ 之间的均方误差作为损失函数

$$\mathcal{L} = \|x_t - \hat{x}_{t|t}\|^2 \quad (12)$$

其中关于 K_t 的梯度可解析计算

$$\frac{\partial \mathcal{L}}{\partial K_t} = 2(K_t \Delta y_t - \Delta x_t) \Delta y_t^\top \quad (13)$$

$$\Delta x_t = x_t - \hat{x}_{t|t-1} \quad (14)$$

该梯度可以通过时间反向传播 (BPTT) 流入 RNN 内部^{[28][29]},实现数据驱动与物理模型的无缝协同优化。

然而,现有 DKF 方法普遍采用 RNN 作为时序骨干,由此带来新的挑战:顺序计算导致推理延迟较高;受梯度消失问题制约,难以捕获超长序列中的缓慢漂移;当观测信号长时间中断后突然恢复时,瞬态冲击易扰乱 RNN 隐状态,引发“恢复期尖峰”现象,严重影响系统稳定性。综上所述,可微卡尔曼滤波通过将数据驱动增益预测嵌入“预测—更新”递归结构,为模型失配和噪声统计未知条件下的状态估计提供了新的实现方式。然而,现有方法主要采用 RNN 对历史新息和状态演化特征进行编码,其顺序递推方式在长序列建模和并行计算方面存在一定限制。为此,本文将具备选择性扫描

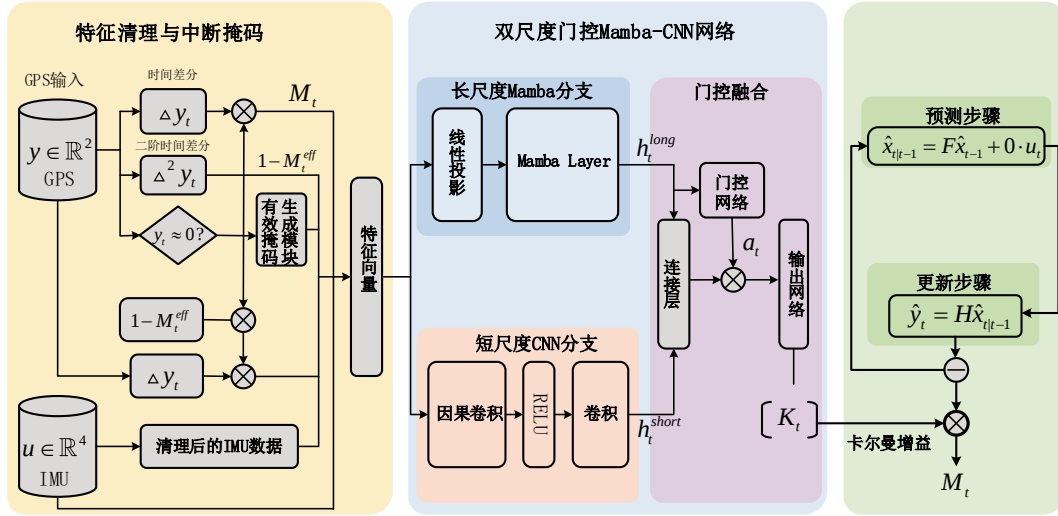


图3 DSG-MC DKF模型结构图

机制的 Mamba 引入 DKF 增益预测网络, 利用输入相关的状态转移和输入映射参数, 对历史 IMU、观测差分、新息及断联状态进行选择编码, 为后续数据驱动增益矩阵的预测提供长程上下文特征。需要说明的是, Mamba 选择性扫描并不直接执行经典卡尔曼增益的解析计算, 其作用位于“历史特征编码—增益预测—新息修正”过程的前端。

3 双尺度门控 Mamba-CNN 可微卡尔曼滤波器(DSG-MC DKF)

本文提出的双尺度门控 Mamba-CNN 可微卡尔曼滤波器 (DSG-MC DKF), 用于突破传统可微卡尔曼滤波在长时序建模与抗干扰能力上的瓶颈, 其整体模型结构如图3所示。该算法在经典卡尔曼滤波的“预测-更新”递归框架内, 通过三大核心创新构建了一个高鲁棒性的端到端融合定位系统。首先, 在训练策略层面, 引入从单窗口到双窗口的演进机制, 有效缓解沿时间步反向传播带来的梯度爆炸与信息衰减问题; 其次, 在网络架构层面, 构建基于 Mamba 与因果 1D-CNN 的双尺度并行特征提取与动态门控融合架构, 实现长程偏置补偿与高频瞬态噪声抑制的有效解耦; 最后, 在核心细节设计层面, 本文采用无显式控制输入的标称状态传播方式, 并引入显式断联掩码与特征清洗机制, 从而降低外部逐时刻姿态旋转结果和异常量测对状态估计过程的直接影响。在量测异常期间, 显式掩码阻断无效 GPS 新息, 状态主要依靠标称状态转移先验进行有限时域传播。

3.1 双窗口架构

在可微卡尔曼滤波的训练过程中, 网络需要隐式学习系统状态的时序演化规律。然而, 传统的沿时间步反向传播 (Backpropagation Through Time, BPTT) 在面对长序列传感器数据时, 极易因连乘效应引发梯度爆炸或梯度消失^{[28][29]}。为此, 常规做法是引入窗口化训练策略, 将总长度为 T 的长序列分割为固定长度 L 的子窗口进行分段优化。但简单的单窗口机制存在固有缺陷: 若窗口设置过短, 模型会彻底丧失对缓慢累积误差 (如 IMU 零偏漂移) 的全局感知能力; 若窗口过长, 不仅再次面临梯度消失风险, 且窗口内的微观时序细节往往被粗暴的全局池化所掩盖, 破坏了原始信号的高频完整性。

为了突破长短期特征相互制约的瓶颈, 本文在特征提取前端将其重构为双窗口并行架构如图4所示。该架构不再对单一输入进行折中处理, 而是并行地构建两种具有不同时间分辨率与感受野的历史信息视图。

具体而言, 定义当前时刻为 t , 长窗口序列 X_{long} 通过一定的步长采样或特征抽象, 覆盖较大的历史时间跨度 $[t - T_L, t]$, 是为了捕获宏观、缓变的物理趋势; 而短窗口序列 X_{short} 则保留原始高分辨率, 仅截取紧邻当前时刻的局部时间片段 $[t - T_S, t]$, 专门用于精确感知瞬态冲击与高频观测噪声。两个窗口的特征流被独立送入后续异构网络中。这种物理意义明确的时空解耦设计, 使模型既能有效补偿累积漂移, 又能快速响应突发动态, 为后续高阶时序建模奠定了坚实的数据组织基础。

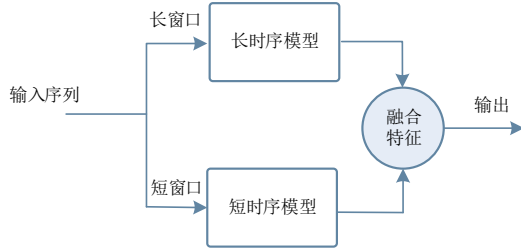


图4 双窗口模型架构图

3.2 Mamba 与 1D-CNN 双尺度并行融合

在确立了双窗口输入机制后，如何针对长短序列的物理特性选择最匹配的网络拓扑便成为关键。近年来，以 Mamba 为代表的结构化状态空间模型（Structured State Space Model, SSM）在序列建模领域展现出较好的潜力^{[30][31]}。SSM 起源于现代控制理论中的连续时间线性时不变系统，其核心是通过隐状态 $h(t)$ 描述输入 $x(t)$ 到输出 $y(t)$ 的演化过程

$$h'(t) = A_s h(t) + B_s x(t) \quad (15)$$

$$y(t) = C_s h(t) \quad (16)$$

其中， $x(t)$ 为输入特征， $h(t)$ 为 SSM 隐状态， $y(t)$

为输出特征； A_s 、 B_s 、 C_s 分别表示 Mamba 状态空间模型中的状态转移矩阵、输入映射矩阵和输出映射矩阵。下标 s 用于区分 Mamba 的状态空间矩阵与卡尔曼滤波中的系统矩阵。

为适应离散化数字信号，经典方法采用线性时不变假设，虽可通过卷积加速，但固定参数的设计严重限制了模型根据输入内容动态筛选信息的能力。Mamba 的核心创新即选择性状态空间模型，将关键参数 B 、 C 及离散化步长 Δ 定义为当前输入 x_t 的动态函数

$$B_{s,t} = \text{Linear}_N(x_t) \quad (17)$$

$$C_{s,t} = \text{Linear}_N(x_t) \quad (18)$$

$$\Delta_t = \text{Softplus}(\text{Linear}_1(x_t)) \quad (19)$$

其中， $B_{s,t}$ 、 $C_{s,t}$ 和 Δ_t 均由当前输入 x_t 动态生成，体现了 Mamba 根据输入内容进行选择性状态更新的机制。

为实现连续系统的精确离散化，Mamba 采用零阶保持规则，计算出时变的离散状态转移矩阵 $\bar{A}_t$ 和输入矩阵 $\bar{B}_t$

$$\bar{A}_t = \exp(\Delta_t A) \quad (20)$$

$$\bar{B}_t = (\Delta_t A)^{-1} (\exp(\Delta_t A) - I) \cdot \Delta_t B_t \quad (21)$$

进而得到具备上下文感知能力的离散递归形式

$$h_t = \bar{A}_t h_{t-1} + \bar{B}_t x_t \quad (22)$$

$$y_t = C_t h_t \quad (23)$$

其中， $\bar{A}_t$ 和 $\bar{B}_t$ 分别表示离散化后的时变状态转移矩阵和输入矩阵，用于将连续时间 SSM 转换为离散递推形式； h_t 为离散时刻的 SSM 隐状态， y_t 为 Mamba 分支输出特征。

得益于硬件感知并行算法与 SRAM 内核融合策略，Mamba 在维持 $O(N)$ 线性时间复杂度的前提下，具备较强的长程依赖建模能力^[31]。为进一步说明 Mamba 状态递推与卡尔曼滤波状态预测之间的关系，以下从递归结构角度进行分析。为避免与卡尔曼滤波中的物理状态 x_t 混淆，将输入 Mamba 分支的时序特征记为 z_t 。由式 (22) 可知，Mamba 的离散隐状态递推为

$$h_t = \bar{A}_t h_{t-1} + \bar{B}_t z_t \quad (24)$$

对该递推关系进行逐步展开，可得

$$h_t = \Phi_M(t, 0) h_0 + \sum_{i=1}^t \Phi_M(t, i) \bar{B}_i z_i \quad (25)$$

其中

$$\Phi_M(t, i) = \bar{A}_t \bar{A}_{t-1} \cdots \bar{A}_{i+1}, \Phi_M(t, t) = I \quad (26)$$

式 (24) 表明，时刻 t 的 Mamba 隐状态由初始隐状态以及历史输入特征经过时变状态转移算子加权累积得到。由于 A_t 和 B_t 随当前输入动态变化，不同时刻的历史特征能够根据输入内容以不同权重保留在当前隐状态中，这为长程时序信息的选择性编码提供了结构基础。

与之对应，卡尔曼滤波的一阶状态预测可表示为

$$\hat{x}_{t|t-1} = F_t \hat{x}_{t-1|t-1} + B_t u_t \quad (27)$$

其中， $\hat{x}_{t|t-1}$ 为预测状态， F_t 为由系统动力学模型确定的状态转移矩阵， $B_t u_t$ 为控制输入项。对比式 (24) 与式 (27) 可以发现，Mamba 与卡尔曼状态预测均采用“上一时刻状态递推与当前输入驱动相结合”的一阶状态空间表达形式。因此，二者在递归结构上具有形式相似性，这也是本文采用 Mamba 编码长程时序信息的理论依据。

需要指出的是，上述相似性仅限于一阶递归结构，并不意味着 Mamba 与卡尔曼滤波在状态含义、概率推断机制和统计最优性方面等价。卡尔曼滤波

中的 $\hat{x}_{t|t-1}$ 表示具有明确物理含义的系统状态，且经典卡尔曼增益由预测误差协方差和量测噪声协方差解析计算：

$$K_t = P_{t|t-1} H_t^T (H_t P_{t|t-1} H_t^T + R_t)^{-1} \quad (28)$$

相比之下，Mamba 中的 h_t 是由训练数据驱动形成的隐特征状态，并不直接对应载体位置、速度等物理状态；同时，Mamba 不显式传播预测误差协方差 $P_{t|t-1}$ ，也不执行由 Q_t 和 R_t 决定的 Riccati 递推。因此，Mamba 并不是卡尔曼状态转移方程或协方差递推过程的等价替代，而是利用其输入相关的递归状态更新机制，对增益预测所需要的长程历史特征进行编码。

在本文模型中，Mamba 长程分支输出的 h_t^{long} 不直接作为导航状态估计结果，而是与短程 CNN 特征、观测新息及断联标志共同用于预测数据驱动增益矩阵。宏观长窗口序列输入后，Mamba 分支有助于提取与缓慢累积偏差和长期运动趋势相关的时序特征，为后续卡尔曼新息更新提供长程上下文信息。

与此同时，针对微观短程分支，我们将覆盖短窗口的高分辨率序列 X_{short} 输入给因果一维卷积网络（1D-CNN）。因果卷积（Causal Convolution）不仅天然擅长提取局部模式，更严格保证了 t 时刻的输出仅依赖于当前及历史时刻的信息，完美契合滤波器的实时性要求^[32]。多层堆叠的 1D-CNN 高效地过滤并提取了短时内的信号突变与高频观测特征，输出 h_t^{short} 。上述基于 Mamba 与 1D-CNN 的双尺度网络内部架构细节如图 5 所示。

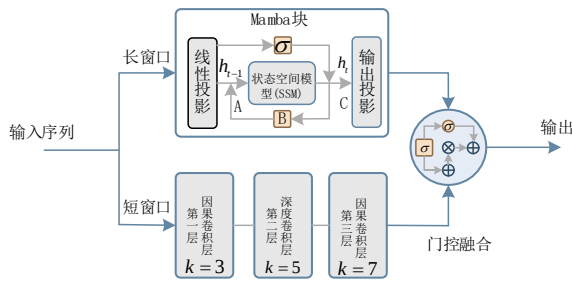


图 5 Mamba 和 1D-CNN 模型结构图

为实现长程时序特征与短程局部特征的自适应融合，两路异构分支的输出被送入由全连接层和 Sigmoid 激活函数构成的动态门控层。结合本文实际使用的门控输入，动态融合权重表示为

$$\alpha_t = \sigma(W_g[h_t^{\text{long}}; h_t^{\text{short}}; \tilde{v}_t; M_t] + b_g) \quad (29)$$

其中， h_t^{long} 和 h_t^{short} 分别为 Mamba 长程分支和 1D-CNN 短程分支的输出特征， $\tilde{v}_t$ 为归一化观测新息， M_t 为 GNSS 断联标志， $\alpha_t \in [0,1]$ 为动态融合权重。两路特征融合方式如下：

$$h_t^{\text{fused}} = \alpha_t \odot h_t^{\text{long}} + (1 - \alpha_t) \odot h_t^{\text{short}} \quad (30)$$

由于 α_t 受到 Sigmoid 函数约束，式（30）构成两路特征之间的有界加权组合，使网络能够根据当前运动状态、观测变化及 GNSS 有效性，自适应调节长程信息和短程信息的相对贡献。

为进一步明确选择性扫描与增益预测之间的联系，由前述 Mamba 递推展开可知， h_t^{long} 由历史输入特征经过输入相关的状态转移算子加权累积得到。因此，Mamba 内部的 $\bar{A}_t$ 、 $\bar{B}_t$ 、 C_t 和离散化步长 Δ_t 并不直接构成卡尔曼增益，而是决定历史 IMU、观测差分、新息和断联信息以何种方式写入、保留并输出为当前长程特征。

融合后的双尺度特征进一步通过增益输出映射转换为向量化增益，并重构为与状态维度和观测维度相匹配的矩阵：

$$k_{\theta,t} = \mathcal{G}_K(h_t^{\text{fused}}) \quad (31)$$

$$K_{\theta,t} = \text{Mat}_{m \times n}(k_{\theta,t}) \quad (32)$$

其中， $\mathcal{G}_K(\cdot)$ 表示由全连接输出层构成的增益映射函数， $K_{\theta,t} \in \mathbb{R}^{m \times n}$ 为向量化增益， $\text{Mat}_{m \times n}(\cdot)$ 表示将其重构为 $m \times n$ 维增益矩阵。观测新息表示为

$$v_t = y_t - H_t \hat{x}_{t|t-1} \quad (33)$$

所预测的增益矩阵最终通过标准新息更新通道参与状态修正：

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + \gamma_t K_{\theta,t} v_t, \gamma_t = 1 - M_t \quad (34)$$

由上述公式可知，Mamba 选择性扫描与卡尔曼增益更新之间形成的是间接的功能联系。首先，选择性扫描根据输入内容调节历史信息在长程隐状态中的写入、保留和输出，形成长程特征 h_t^{long} ；其次，动态门控根据长程特征、短程特征、归一化新息及断联状态生成融合权重 α_t ，得到双尺度融合特征 h_t^{fused} ；随后，增益输出层将融合特征映射为 $K_{\theta,t}$ ；最后， $K_{\theta,t} v_t$ 决定当前观测新息对预测状态的修正强度和方向。

因此，Mamba 选择性扫描并不直接计算经典

卡尔曼增益,也不等价于误差协方差或 Riccati 递推。经典卡尔曼增益由 $P_{t|t-1}$ 、 H_t 和 R_t 解析计算,而本文的 $K_{\theta,t}$ 由历史时序特征经过网络映射得到。二者均用于调节预测信息与观测信息之间的相对贡献,但其信息来源和计算方式不同。Mamba 选择性扫描的作用是构建增益预测所需要的历史上下文表示,增益输出层则在训练数据和轨迹损失约束下形成任务相关的增益近似。

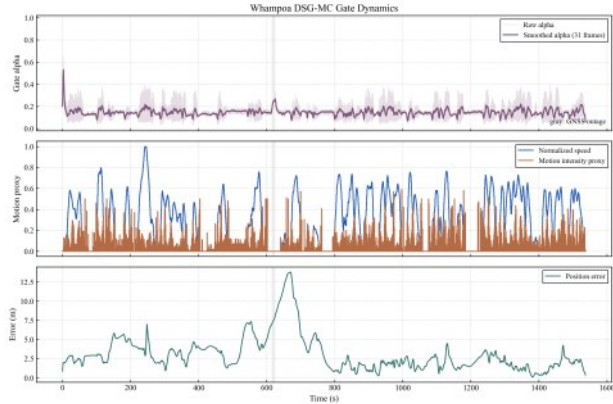


图 6 动态门控在不同运动模式下的实际调节行为

动态门控权重并非人工设定,而是由当前时刻的长程 Mamba 特征、短程 CNN 特征、归一化观测新息以及 GNSS 断联标志共同输入门控网络后自适应生成。由于门控输出经过 Sigmoid 函数约束,其取值被限制在 $[0,1]$ 范围内,因此融合过程可视为长程趋势特征与短程瞬态特征之间的有界凸组合。在平稳直行或 GNSS 断联等需要抑制短时噪声的场景下,门控更倾向于保留 Mamba 分支中的长程运动记忆;在转弯、加减速或观测新息快速变化的场景下,门控可提高短程 CNN 分支的贡献,从而增强模型对局部动态变化的响应能力。最终门控权重通过端到端轨迹误差反向传播进行学习,无需手工阈值调节。

为进一步验证双分支动态门控在不同运动模式下的实际调节行为,本文统计了 Whampoa 路线中融合权重 α_t 的时间变化,如图 6 所示。结果表明,在 GNSS 断联区间,长程 Mamba 分支权重由正常阶段的 0.147 提升至 0.234;而在高转弯和高加速度阶段, α_t 略有下降,短程 CNN 分支贡献相应增强。该现象表明,双分支融合门控能够根据断联状态和运动变化强度,在长程趋势特征与短程动态特征之间进行自适应权衡。

本文统计的 α_t 属于 Mamba 分支与 CNN 分支之间的特征融合门控,并不等同于 Mamba 内部选择性扫描所采用的 A_t 、 B_t 、 C_t 及离散化步长 Δ_t 。二者分别作用于不同层级: Mamba 选择性扫描用于调节历史信息在长程隐状态中的写入、保留与输出方式,双分支动态门控则用于调节长程特征和短程特征在增益预测过程中的贡献比例。两类机制最终均通过融合特征影响 $K_{\theta,t}$,但不存在参数上的直接对应关系。

3.3 隐式动力学建模与显式断联掩码

作为 DSG-MC DKF 的核心设计之一,本节进一步阐述隐式动力学建模与显式时序掩码机制。二者分别从状态演化建模和异常量测抑制两个层面增强模型在复杂导航场景下的鲁棒性,并为后续卡尔曼更新过程提供具有物理约束的残差修正通道。这两大机制针对真实复杂环境进行了自底向上的设计重构,共同赋予模型端到端学习能力与极端工况下的鲁棒性。在传统 IMU/GNSS 融合定位中,预测步骤通常根据外部姿态估计结果,将载体坐标系下的 IMU 量测转换至导航坐标系,并通过离散运动学方程形成显式控制输入。该过程具有明确的物理意义,但其精度会受到姿态解算误差、传感器安装偏差、时间同步误差及运动学模型失配的共同影响。为减少上述误差通过显式控制输入直接传播至状态预测过程,本文将控制输入矩阵置零,即 $B = 0$,使预测通路由状态转移矩阵 F 完成标称状态先验传播。

$$\hat{x}_{t|t-1} = F\hat{x}_{t-1|t-1} + 0 \cdot u_t \quad (35)$$

其中, F 为状态转移矩阵, B 为控制输入矩阵, u_t 为 IMU 输入。当 $B = 0$ 时,显式控制输入项 B 被置零,预测步退化为仅由 F 控制的状态先验传播。

该设计并不意味着舍弃 IMU 信息,而是取消其在预测方程中的显式注入方式。换言之,IMU 信息不再通过固定的 $B_t u_t$ 项直接作用于状态预测,而是与观测新息、差分特征及断联掩码共同输入到双尺度 Mamba-CNN 网络中,由网络在数据驱动的更新通道中隐式学习其对状态修正的影响。进一步地,从模型失配角度看,真实复杂导航系统的状态演化可表示为标称状态转移、显式控制输入以及多源未建模误差的叠加,即

$$x_t = F_t x_{t-1} + B_t u_t + \psi_t^{\text{coord}} + G_b b_t + d_t + w_t \quad (36)$$

其中, F_t 和 B_t 分别为标称状态转移矩阵和控制输入

矩阵, u_t 为 IMU 量测输入; ψ_t^{coord} 表示局部载体坐标系与全局坐标系之间的姿态、安装误差及坐标转换近似所引入的模型失配; b_t 表示 IMU 缓慢变化的零偏状态, G_b 为零偏误差到系统状态空间的映射矩阵; d_t 表示未被标称模型描述的机动、振动及环境扰动; w_t 表示随机过程噪声。为便于后续分析, 将上述未由标称状态转移模型准确描述的误差统一记为

$$\psi_t = B_t u_t + \psi_t^{\text{coord}} + G_b b_t + d_t + w_t \quad (37)$$

需要指出的是, 式 (36) 中的各误差分量具有不同的物理来源, 但它们在观测空间中可能产生相互耦合或相似的误差表现。在仅利用位置观测和轨迹监督的条件下, 各误差分量通常不具备唯一可辨识度。

本文不利用神经网络直接无约束回归系统状态, 也不要求网络分别估计式 (36) 中的各类误差分量。所提方法首先利用标称状态转移模型完成状态预测:

$$\hat{x}_{t|t-1} = F_t \hat{x}_{t-1|t-1} \quad (38)$$

并根据预测状态构造观测新息:

$$v_t = y_t - H_t \hat{x}_{t|t-1} \quad (39)$$

双尺度 Mamba-CNN 网络以历史 IMU 序列、观测差分、新息特征及断联标志构成的特征序列 Z_t 为输入, 预测数据驱动增益矩阵:

$$K_{\theta,t} = \mathcal{G}_{\theta}(Z_t) \quad (40)$$

最终状态更新表示为

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + \gamma_t K_{\theta,t} v_t, \gamma_t = 1 - M_t \quad (41)$$

由式 (38) - 式 (41) 可知, 网络输出并不直接对应式 (36) 中的坐标映射误差、IMU 零偏或未建模扰动, 也不直接替代完整的状态动力学方程。网络实际学习的是历史传感器特征和观测新息到增益矩阵 $K_{\theta,t}$ 之间的任务相关映射, 并通过 $K_{\theta,t} v_t$ 调节观测信息对状态估计的修正作用。因此, 模型所实现的是受标称状态转移先验、观测模型、新息方向和量测有效性门控共同约束的综合误差补偿, 而不是对各类物理误差源进行无约束、唯一的分量辨识。

为进一步说明综合模型失配与网络修正之间的关系, 定义预测误差和更新误差分别为

$$e_{t|t-1} = x_t - \hat{x}_{t|t-1}, e_{t|t} = x_t - \hat{x}_{t|t} \quad (42)$$

将式 (36) 和式 (38) 代入预测误差定义, 可得

$$e_{t|t-1} = F_t e_{t-1|t-1} + \psi_t \quad (43)$$

其中, ψ_t 为式 (37) 定义的综合模型失配项。若观测方程表示为

$$y_t = H_t x_t + v_t \quad (44)$$

则观测新息可进一步写为

$$v_t = H_t e_{t|t-1} + v_t \quad (45)$$

将式 (41) 和式 (45) 代入更新误差定义, 可得

$$e_{t|t} = (I - \gamma_t K_{\theta,t} H_t) e_{t|t-1} - \gamma_t K_{\theta,t} v_t \quad (46)$$

式 (43) 表明, 坐标映射误差、IMU 零偏、未建模运动和过程噪声首先以综合模型失配 ψ_t 的形式影响预测误差; 式 (46) 则表明, 网络只能通过调节增益矩阵 $K_{\theta,t}$, 改变观测新息对预测误差的修正强度和方向。因此, 网络对综合模型失配的补偿能力受到观测矩阵 H_t 、新息可用性、增益表达能力以及训练数据覆盖范围的共同限制, 而不能无条件地恢复或消除任意未建模误差。

从可学习性角度看, 本文模型存在以下约束。首先, 当不同误差源在观测空间中产生近似相同的新息响应时, 仅依靠轨迹监督难以对坐标映射误差、IMU 零偏和未建模扰动进行唯一分解, 网络只能学习与状态估计目标相关的等效补偿关系。其次, 网络修正项具有 $K_{\theta,t} v_t$ 的形式, 当某一状态误差位于观测矩阵 H_t 的不可观测子空间, 或观测新息不能充分反映该误差时, 网络无法仅依靠当前量测更新保证该误差得到准确修正。最后, 网络参数是在有限训练数据分布上通过轨迹损失优化得到的, 其所学习的综合误差补偿规律主要适用于训练样本所覆盖的运动状态、传感器偏置和噪声范围。由此, 本文所称的“隐式动力学建模”应理解为在给定数据和观测约束下, 对历史传感器特征到新息修正增益之间的任务相关映射进行学习, 而不是对真实完整动力学方程进行无条件辨识。

然而, 仅靠隐式演化不足以应对极端的量测异常。当 GPS 等绝对观测信号长时间丢失或骤然恢复时, 原始观测值新息的一阶差分和二阶差分会产生显著异常尖峰。如果网络被动接收这些“脏数据”, 状态预测将瞬间发散。为此, 本文设计了基于时序偏移逻辑的显式掩码清洗机制。

定义原始断联指示器 $M_t \in \{0,1\}$, 为彻底规避信号恢复瞬间的恢复期尖峰, 构建包含一帧历史记忆的有效掩码 M_t^{eff}

$$M_t^{\text{eff}} = \min(M_t + M_{t-1}, 1) \quad (47)$$

其中, $M_t \in \{0,1\}$ 为原始 GPS 断联标志, 当 GPS 观测缺失或无效时 $M_t = 1$, 否则 $M_t = 0$; M_t^{eff} 为包含一帧历史记忆的有效断联掩码, 当前帧或上一帧存在断联时, $M_t^{\text{eff}} = 1$, 用于屏蔽断联及恢复瞬间产生的异常差分特征。

这意味着只要当前帧或上一帧存在断联, 系统就会判定当前时刻的差分特征失效。基于此掩码, 模型对观测特征执行置零与限幅处理, 生成高度鲁棒的网络输入特征空间 Z_t

$$\Delta^2 y_t = \Delta y_t - \Delta y_{t-1} \quad (48)$$

$$Z_t = [(1 - M_t^{\text{eff}})\Delta y_t, (1 - M_t^{\text{eff}})\Delta^2 y_t, \text{Clamp}(u_t), M_t]^T \quad (49)$$

其中, Δy_t 表示 GPS 观测一阶差分, $\Delta^2 y_t$ 表示 GPS 观测二阶差分, 二者分别用于描述局部运动变化和短时观测突变; Z_t 表示输入双尺度 Mamba-CNN 网络的特征向量, 由清洗后的一阶差分、二阶差分、IMU 输入以及断联标志共同构成。

进一步地, GPS 拒止期间的滤波一致性可从量测门控更新角度进行说明。当 GPS 正常时, $M_t = 0$, 此时 $\gamma_t = 1$, 滤波器根据观测新息进行状态修正; 当 GPS 处于断联状态时, $M_t = 1$, 此时 $\gamma_t = 0$, 状态更新退化为

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} \quad (50)$$

这表明在无效量测阶段, 滤波器不会利用缺失或异常 GPS 观测执行 measurement update, 从而避免虚假新息污染状态估计。此时估计误差主要由状态转移先验和未建模扰动传播, 可写为

$$e_{t|t} = F e_{t-1|t-1} + \delta_t \quad (51)$$

其中, $e_{t|t} = x_t - \hat{x}_{t|t}$, δ_t 表示过程噪声、IMU 残余漂移和未建模动态共同形成的综合误差。对于长度为 L 的有限拒止区间, 有

$$\|e_{t+L}\| \leq \|F\|^L \|e_t\| + \sum_{i=0}^{L-1} \|F\|^i \|\delta_{t+L-1-i}\| \quad (52)$$

式 (50) - 式 (52) 表明, 显式量测门控能够阻断缺失或异常 GPS 新息对状态更新的直接污染, 但并不能消除标称状态传播误差、IMU 残余漂移和未建模扰动在拒止期间的持续累积。为进一步明确

有限拒止区间内的误差边界, 假设

$$\|F\| \leq \rho, \quad \|\delta_t\| \leq \bar{\delta} \quad (53)$$

则在 GNSS 拒止区间内, 状态误差满足

$$\|e_{t+L}\| \leq \rho^L \|e_t\| + \bar{\delta} \sum_{i=0}^{L-1} \rho^i \quad (54)$$

进一步可写为

$$\|e_{t+L}\| \leq \begin{cases} \rho^L \|e_t\| + \bar{\delta} \frac{1 - \rho^L}{1 - \rho}, & 0 \leq \rho < 1, \\ [8pt] \|e_t\| + L\bar{\delta}, & \rho = 1, \\ [6pt] \rho^L \|e_t\| + \bar{\delta} \frac{\rho^L - 1}{\rho - 1}, & \rho > 1. \end{cases} \quad (55)$$

上述结果说明, 当 $\rho < 1$ 时, 有限扰动条件下的误差可保持有界; 当 $\rho = 1$ 时, 误差上界可能随断联长度近似线性增长; 当 $\rho > 1$ 时, 误差上界可能随断联长度快速增大。因此, 显式断联掩码所提供的是对异常量测更新的隔离能力, 而不是对任意时长 GNSS 拒止的无条件稳定性保证。本文关于长时间拒止鲁棒性的结论仅适用于实验所覆盖的有限断联长度和扰动范围。

此外, 有效掩码 M_t^{eff} 进一步屏蔽了断联恢复后一帧产生的异常差分尖峰, 使滤波器能够在恢复阶段更加平滑地重新引入观测修正, 但其效果仍取决于恢复后观测质量、预测误差大小及网络训练分布。

在此过程中, M_t 被作为独立的维度显式拼接至输入特征中。这并非简单的数据预处理, 而是向神经网络注入了一条明确的物理指令。当遭遇信号盲区即 $M_t = 1$ 时, 这条指令会触发前文所述的动态门控机制, 引导网络将决策权重全面倒向 Mamba 长程分支。此时, Mamba 模型切断外部观测输入, 使模型主要依赖状态转移先验以及 Mamba 分支中编码的历史惯性运动特征维持轨迹连续性; 而当信号重新恢复且跨越尖峰期后, 门控层又迅速唤醒 1D-CNN 分支, 重新接管高频新息的修正工作。该显式掩码清洗机制的完整工作原理与对应波形如图 7 所示。

需要说明的是, 所提网络的学习能力受到训练数据、历史感受野、观测信息和增益性质的共同约束。网络学习的是训练分布内历史传感器特征到数据驱动增益之间的经验映射, 当运动状态、传感器噪声、IMU 零偏或 GNSS 遮挡模式显著超出训练范

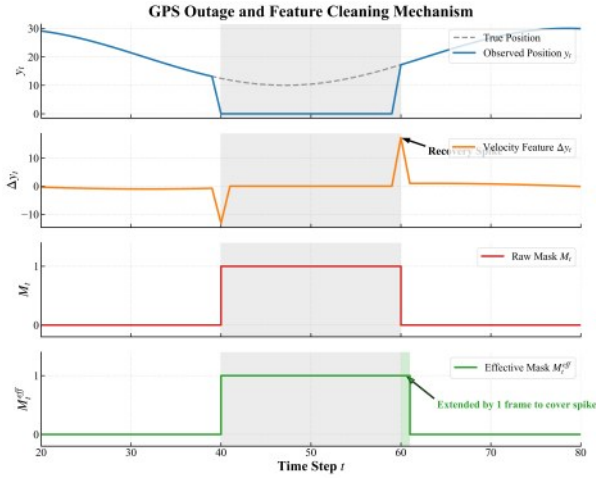


图7 显式掩码清洗机制原理图

围时，其增益预测性能可能下降。Mamba 长程分支能够利用的历史信息也受到输入窗口长度、隐状态维度和训练序列长度限制，并不意味着网络能够无损记忆任意长时间跨度的信息。此外，由于状态修正采用 $\gamma_t K_{\theta,t} v_t$ 的形式，网络不能突破观测模型和新息可用性所决定的信息边界。所预测的 $K_{\theta,t}$ 由轨迹损失监督获得，并非通过显式协方差递推解析计算，因此不自动具有经典卡尔曼增益的统计最优性，也不能为任意动力学、噪声条件和 GNSS 拒止长度提供全局稳定性保证。

综上所述，本文所提出的 DSG-MC DKF 框架并非以黑盒网络直接替代完整滤波过程，而是在经典卡尔曼“预测—更新”递归结构内引入数据驱动的增益预测机制。一方面，状态预测仍由标称状态转移矩阵 F_t 维持基本运动先验；另一方面，历史 IMU 序列、观测差分、新息特征和断联标志共同输入双尺度 Mamba-CNN 网络，用于预测与当前时序状态相匹配的增益矩阵 $K_{\theta,t}$ ，并通过 $K_{\theta,t} v_t$ 调节观测新息对预测状态的修正作用。该过程能够在训练数据覆盖范围内缓解坐标转换近似、IMU 零偏及未建模运动等多源模型失配对状态估计的综合影响，但并不意味着网络能够分别辨识各类误差分量，或无条件恢复真实完整动力学。与此同时，显式断联掩码通过量测门控和特征清洗机制抑制无效 GPS 观测及恢复期尖峰对状态更新的干扰，双尺度门控网络则在长程惯性趋势和短程观测扰动之间进行自适应权衡。由此，本文方法在保留经典滤波递归结构和物理可解释性的同时，增强了模型对长时漂移、突发断联和观测恢复冲击的鲁棒处理能力。需要说

明的是，由于卡尔曼增益 $K_{\theta}(Z_t)$ 仍由训练数据驱动的网络生成，因此该隐式残差修正能力与训练样本所覆盖的运动模式、传感器噪声和 GNSS 退化模式有关；当测试场景显著超出训练分布时，模型性能仍可能受到影响。

4 实验与结果分析

为全方位评估所提及的双尺度门控 Mamba-CNN 可微卡尔曼滤波器 (DSG-MC DKF) 在各类极端非线性情况及真实自动驾驶场景下的鲁棒性能，本文设计了由纯理论仿真演进至真实场景零样本 (Zero-shot) 部署的四组递进式验证实验，来全面解析架构特性并确立其在复杂导航任务中的适用性。同时 Revach 等人提出的 KalmanNet 是 DKF 领域最具代表性的工作^[27]。该方法以新息序列为输入，利用 GRU 网络动态预测卡尔曼增益，在雷达跟踪、机器人定位等任务中显著优于 EKF。因此本文选择 KalmanNet 和经典卡尔曼滤波来作为对比算法。实验首先在动力学先验失配、强非线性混沌系统和缓变漂移误差场景下验证所提模型的基础状态估计能力；随后在真实 UrbanNav 城市峡谷数据集上进一步评估其对真实 GNSS 噪声、遮挡和观测异常的适应能力。各实验根据任务特点设置相应对比算法，其中标准卡尔曼滤波、GRU DKF 及相关 DSG-MC 变体作为主要状态估计基线；在真实数据实验中，本文进一步引入典型深度时序预测模型进行扩展对比。

4.1 实验设置与实现细节

实验数据采样频率为 10 Hz，训练序列长度设为 $T = 200$ 。状态向量维度设为 $m = 4$ ，包括二维位置与二维速度，即 $[x_t, y_t, v_x, v_y]$ 。观测向量维度设为 $n = 2$ ，对应 GPS 平面位置观测 $[x_t, y_t]$ 。

在输入特征构造方面，网络输入由观测差分、二阶差分、IMU 输入以及 GPS 断联标志共同组成。具体而言，首先根据连续 GPS 观测构造位置差分特征，用于近似描述相邻时刻的局部运动变化；进一步对该差分项再次求差，得到二阶差分特征，用于表征观测突变、恢复尖峰及短时扰动信息。为避免 GPS 断联和信号恢复瞬间产生异常尖峰，当前帧或前一帧处于断联状态时，对差分特征进行清洗。随后，IMU 输入经过幅值裁剪后，与 GPS 差分特征、二阶差分特征及断联 mask 拼接，形成最

终网络输入。

在双窗口时序建模方面, DSG-MC DKF 采用长短时序分支协同建模策略。其中, 长程分支采用 1 层 Mamba 模块, 其隐藏维度为 $d_{model} = 64$, 状态维度为 $d_{state} = 16$, 卷积宽度为 $d_{conv} = 4$, 扩展倍率为 $d_{expand} = 2$, 长窗口长度设为 $w_l = 40$, 用于捕获较长时间范围内的运动趋势、惯性信息累积以及误差演化规律; 短程分支采用两层 1D-CNN, 第一层卷积核大小为 5, 第二层卷积核大小为 3, 短窗口长度设为 $w_s = 5$, 用于提取局部动态变化、观测跳变和短时高频扰动特征。两类时序特征经过动态门控融合模块进行自适应加权, 从而使模型能够在平稳运动、GPS 断联以及信号恢复阶段动态调整长短期信息的贡献。训练过程中采用 Adam 优化器, Mamba 系列模型学习率设为 $l = 5 \times 10^{-4}$, 训练轮数为 300, batch size 设为 512, 损失函数采用状态估计均方误差 MSE。

4.2 线性动力学先验失配环境下的状态估计

为验证滤波器在显著动力学先验失配条件下的鲁棒性与自我校正能力, 本文设计了一组基于经典二维线性追踪任务的实验。滤波器内部预设的运动模型保持为理想的线性标称动力学矩阵, 而数据生成引擎中则特意引入了系统性偏差, 在底层真实状态转移过程中施加了 20° 的全局旋转失配角, 用以模拟真实物理环境与理想数学模型之间可能存在的严重非线性偏差, 从而评估算法在动力学失配下的跟踪精度与收敛特性。

定量评估与实验结果表明, 高度依赖标称矩阵的标准卡尔曼滤波器因无法感知底层的旋转畸变, 其状态估计均方误差显著上升。GRU 模型虽然在一定程度上通过可微卡尔曼滤波的结构克服了底层的旋转畸变但是结果仍是不太理想。但是 DSG-MC Gated 模型成功在隐空间内逆向解算出了系统的仿射变换规律, 将误差大幅压低至 -1.5 dB, 初步证实了 SSM 机制在补偿全局系统误差方面的明显优势。各算法在失配环境下的具体状态追踪曲线与绝对误差对比如图 8 所示。

4.3 高阶非线性洛伦兹混沌系统重构

为考察算法在强非线性环境中的适应性, 本文设计了一组基于三维洛伦兹 (Lorenz) 混沌吸引子系统的实验^[33]。鉴于该动力学系统对初始状态具有极端敏感性与高度非线性特征, 其已成为评估状

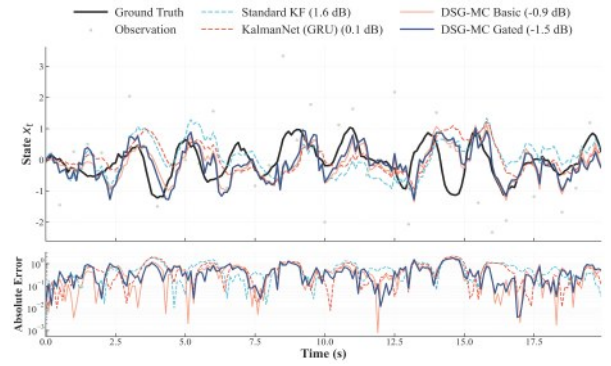


图 8 线性动力学先验失配环境下的状态估计研究结果图

态估计算法处理强非线性动态能力的严苛基准。在此设定下, 本文将算法的先验约束严格限定为单位转

移矩阵, 即在缺失先验动态信息的前提下, 考察各算法从含噪观测数据中重构混沌动力学演化规律的有效性, 通过这一挑战性任务量化评估本文架构对高维非线性特征的重构精度。

真实系统状态在三维相空间内呈现典型的双翼蝴蝶状吸引子结构。在缺少准确非线性动力学模型和雅可比矩阵的条件下, 传统卡尔曼滤波难以有效描述系统的强非线性演化, 估计轨迹与真实吸引子之间出现明显偏离; GRU 网络虽然能够捕捉部分局部变化趋势, 但在吸引子两翼切换区域仍产生较为明显的形态失真。相比之下, Mamba 架构的估计轨迹能够较好保持洛伦兹吸引子的整体几何结构, 并在多数时间段内跟随真实状态的演化趋势。该结果表明, Mamba 的递归隐状态对本文实验条件下的长程非线性时序依赖具有较强的表征能力, 能够为数据驱动增益预测提供有效的历史上下文信息。需要指出的是, 该实验验证的是模型在给定训练和测试分布下的状态重构能力, 并不意味着 Mamba 已经唯一辨识或严格恢复了洛伦兹系统真实的非线性微分算子。各模型对该混沌系统的三维空间重构轨迹对比如图 9 所示。

4.4 应对长时累积误差的零偏自适应补偿

传感器零偏 (Bias) 漂移是制约惯性导航系统长程积分稳定性的关键技术瓶颈。为评估各算法对该类误差的补偿能力, 本文基于 NCLT 数据集的统计特性, 构建了一组跨度长达 400 个时间步的长时序轨迹推演任务^[34]。在该实验设置中, 系统输入的测速数据被叠加了不可观测的随机游走偏差 (Random Walk Bias), 旨在考察算法在复杂噪声干

扰环境下,对缓变漂移误差进行在线感知与有效抑制的鲁棒性能。

结合图 10 局部放大视角下的轨迹重构结果可清晰观察到,传统卡尔曼滤波框架由于缺乏针对零偏误差的在线自适应估计机制,其状态估计轨迹呈现出显著的非线性发散趋势。例如,在实验数据中,随着时间推移,传统卡尔曼滤波的估计误差持续累积,与真实轨迹之间的偏差明显扩大,反映出其对缓变偏置的抑制能力不足。相比之下,本文所提的 Mamba 架构能够利用长窗口内的历史 IMU、观测差分及状态修正信息,提取与缓慢偏置变化相关的长程时序特征,并据此动态调整后续滤波迭代中的增益矩阵,使估计结果对持续累积的偏置误差表现出更强的抑制能力。需要说明的是,该实验结果表明网络学习到了有利于减小零偏影响的等效补偿规律,并不意味着其隐状态已经对真实 IMU 零偏进行了唯一、显式的参数辨识。进一步从图 10 中的数据可以看出,所提方法在轨迹重构中始终紧密跟随真实轨迹,在多个关键时间节点上的估计精度均优于传统卡尔曼滤波、基于 GRU 的 Kalman-

Net 以及 DSG-MC Basic,尤其在长时段内保持了良好的收敛特性。该实验观测结果有力验证

了“长程特征分支负责宏观累积偏置补偿”的架构设计初衷,证明了所提方法在长程时序依赖建模与缓变误差抑制方面的优异性能。

4.5 基于真实香港城市峡谷数据集的零样本泛化测试与适用范围分析

数据驱动增益预测的泛化能力与训练数据所覆盖的运动模式、传感器误差和量测退化条件密切相关。本文所称的零样本测试,是指 UrbanNav-HK TST 和 Whampoa 真实路线未参与网络参数训练,模型在不使用目标路线进行微调的情况下直接完成状态估计。该设置用于评价模型从仿真训练分布向未见真实路线迁移的能力。

为提高仿真训练数据的覆盖范围,训练阶段对轨迹形态、运动速度、加速度、转弯模式、传感器噪声、IMU 零偏漂移及 GNSS 拒止区间进行随机化设置,使网络能够接触多种运动与量测退化组合^[35]。然而,这种域随机化只能扩大训练分布的覆盖范围,不能从理论上消除仿真数据与真实场景

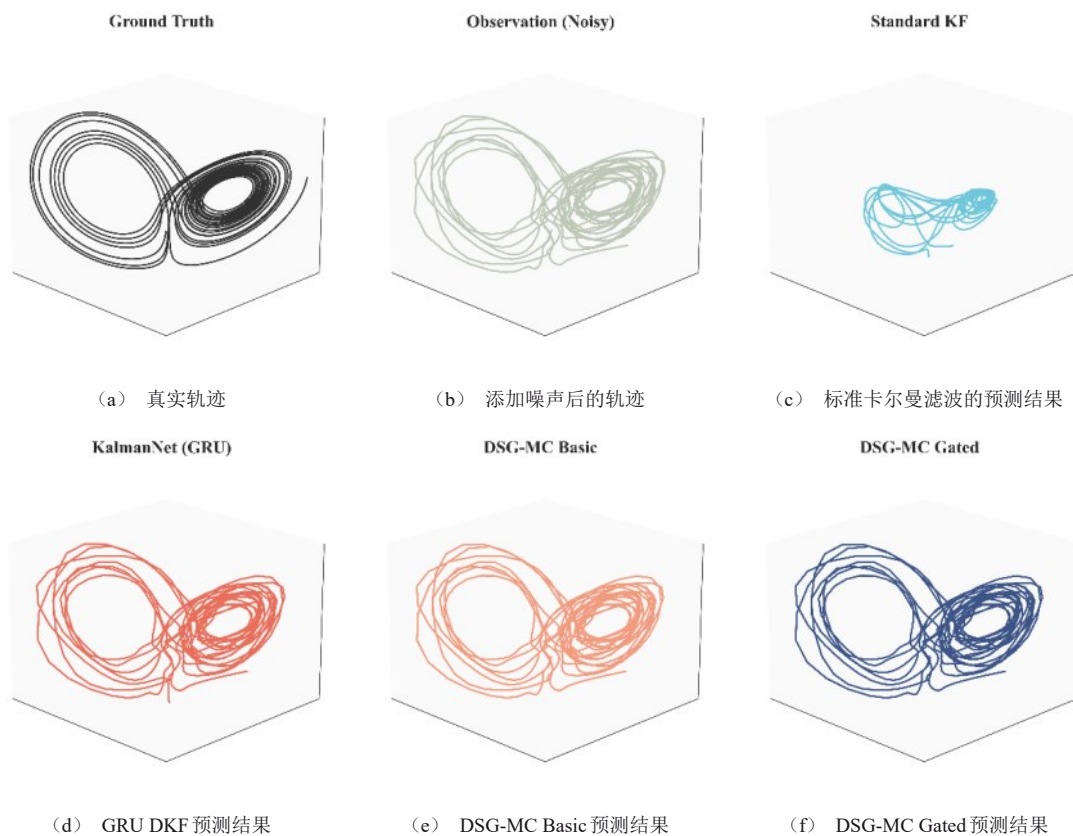


图 9 各模型高阶非线性洛伦兹混沌系统重构实验结果

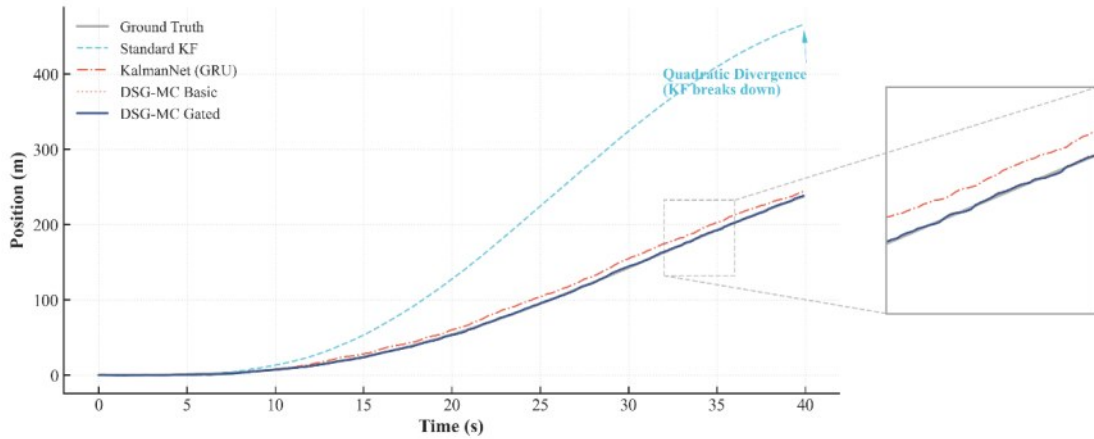


图 10 各模型在缓变漂移误差实验结果

之间的分布差异。当测试场景中的速度范围、转弯半径、加速度分布、IMU 偏置强度、噪声类型或 GNSS 遮挡持续时间显著超出训练阶段的随机化范围时, 输入特征可能偏离网络已学习的有效区域, 进而导致增益预测偏差和状态估计性能下降。

在此基础上, 本文进一步选取 UrbanNav-HK TST 与 Whampoa 两条具有代表性的真实城市峡谷路线开展测试, 并与标准 KF、GRU-DKF、TFT、FEDformer、xLSTM、TabPFN-TS 以及相关 DSG-MC 变体进行对比^[38]。其中, xLSTM 代表采用扩展循环记忆结构的长程时序模型^[36], TabPFN-TS 代表基于预训练表格基础模型的时间序列预测方法^[37]。通过上述对比, 进一步评价不同长程时序建模范式在真实 GNSS 噪声、城市峡谷遮挡和观测恢复异常条件下的状态估计性能。

由图 11 和图 12 可见, 在 UrbanNav-HK TST 路线中, 标准 KF 与 GRU-DKF 均出现较为明显的轨迹偏移, RMSE 分别达到 21.62 m 和 22.27 m, 说明固定运动学模型或单一传统循环结构难以稳定应对真实城市环境中的非平稳量测噪声和局部观测异常。TabPFN-TS 能够恢复轨迹的主要几何轮廓, 但在路线左侧及底部转弯区域仍存在较明显的累计偏差, 其 RMSE 为 10.19 m, 表明将多传感器融合定位直接转化为表格时间序列回归后, 模型性能仍容易受到量测噪声和序列分布变化的影响。

FEDformer、xLSTM 和 TFT 的 RMSE 分别为 4.77、3.06 和 2.83 m。其中, FEDformer 能够拟合轨迹的整体趋势, 但局部转弯区域仍存在一定偏移; xLSTM 利用扩展循环记忆结构改善了长程信息保持能力, 其估计轨迹与真实轨迹具有较高的一

致性; TFT 通过变量选择和门控注意力机制进一步降低了定位误差。DSG-MC Basic 取得 3.11 m 的 RMSE, 说明双尺度 Mamba-CNN 特征提取本身能够改善长序列状态估计, 但在缺少完整动态门控机制时仍存在一定性能限制。

本文提出的 DSG-MC Gated 取得 2.78 m 的最低 RMSE。与 xLSTM、DSG-MC Basic、FEDformer 和 TabPFN-TS 相比, 其 RMSE 分别降低约 9.2%、10.6%、41.7% 和 72.7%; 与 TFT 相比降低约 1.8%。需要指出的是, DSG-MC Gated 与 TFT 之间的数值差距相对较小, 因此该结果不宜表述为显著优于 TFT。更准确地说, 所提模型在获得较高轨迹估计精度的同时, 还保留了标称状态传播、观测新息更新和 GNSS 断联门控等滤波结构, 在数据驱动建模能力与状态估计可解释性之间取得了较好的平衡。

为进一步考察模型在更复杂真实场景下的稳定性, 本文选取 UrbanNav Whampoa 路线进行测试。该路线包含更明显的城市峡谷遮挡、道路转向和 GNSS 局部异常, 对滤波器的抗拒止能力和轨迹拓扑保持能力提出了更高要求。各模型在 Whampoa 路线上的轨迹估计结果如图 13 所示, 定量对比如图 14 所示。

从 Whampoa 路线结果可以看出, 标准 KF 与 GRU-DKF 在复杂城市峡谷场景下均出现明显的轨迹畸变, RMSE 分别达到 42.41 m 和 45.25 m, 说明传统模型驱动方法和单一循环增益网络难以稳定应对该路线中频繁转向、城市峡谷遮挡及非平稳观测异常。xLSTM 能够较好恢复路线的主体几何结构, 并将 RMSE 降低至 3.16 m, 说明扩展循环记忆结构有助于提取长时间范围内的运动趋势, 但在多次急

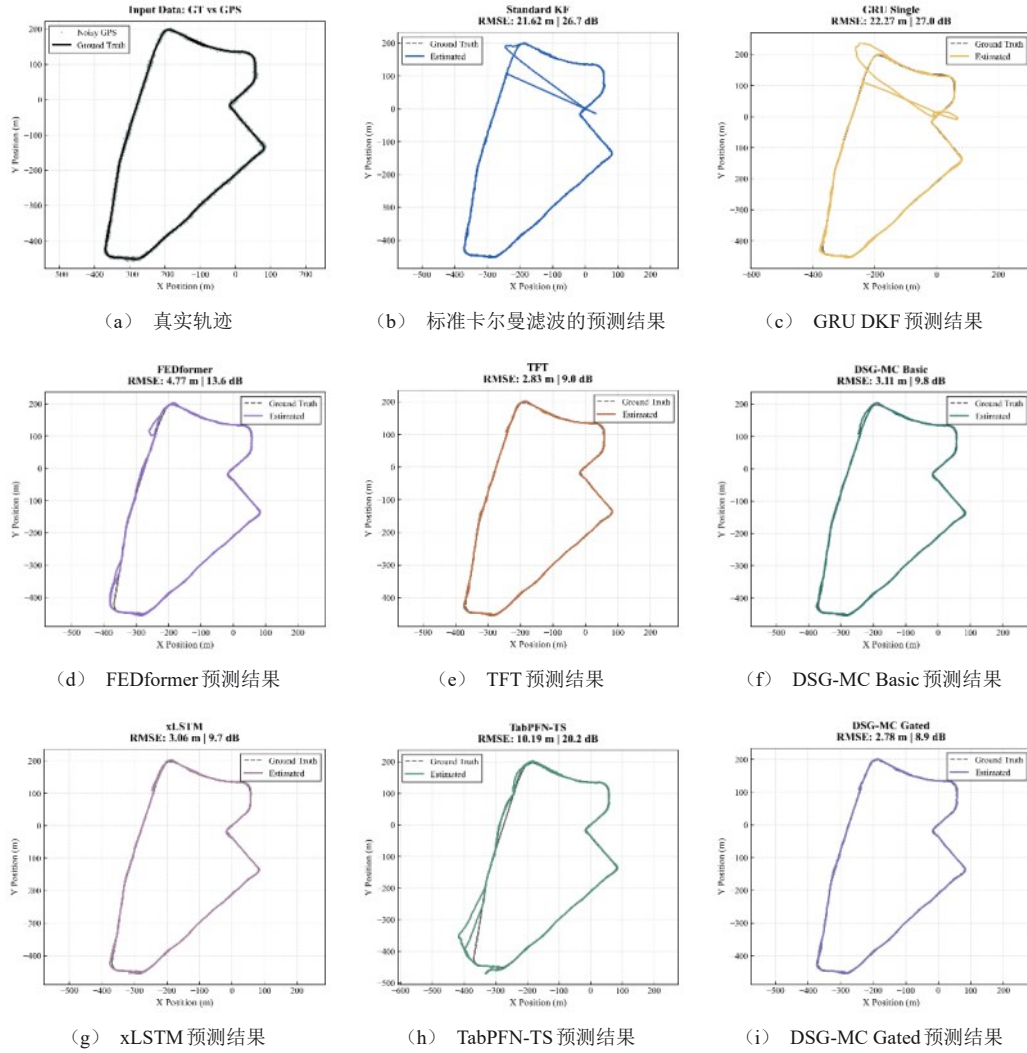


图 11 各模型在 UrbanNav-HK 路线下的轨迹估计结果

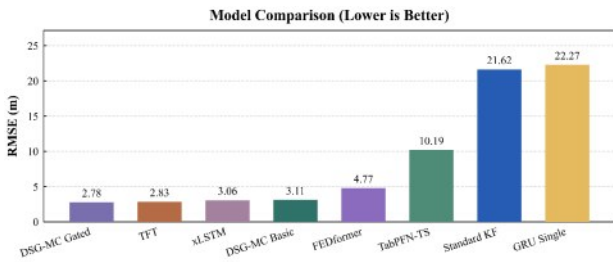


图 12 各个模型在 UrbanNav-HK 数据下的实验结果

转弯和局部观测异常区域仍存在一定累计偏差。

FEDformer、TFT 和 TabPFN-TS 在 Whampoa 路线中的 RMSE 分别为 2.75、2.78 和 2.78 m，均能够较准确地拟合真实轨迹的整体形态。值得注意的是，TabPFN-TS 在 Whampoa 路线中的结果明显优于其在 UrbanNav-HK TST 路线中的 10.19 m，说明该预训练表格时间序列方法对目标路线特征、上下

文样本和量测分布具有较强敏感性，不同路线间的性能一致性仍存在一定局限。

DSG-MC Gated 在 Whampoa 路线中取得 2.74 m 的最低 RMSE，相较于 xLSTM 降低约 13.3%，同时分别比 FEDformer、TFT 和 TabPFN-TS 降低约 0.4%、1.4% 和 1.4%。由于 DSG-MC Gated 与 FEDformer、TFT 和 TabPFN-TS 之间的数值差距较小，因此该结果不宜表述为显著优于上述模型。其主要优势在于，在保持较高定位精度的同时，仍保留了卡尔曼预测—更新递归结构、观测新息修正和显式断联量测门控。

相比之下，DSG-MC Basic 在 Whampoa 路线中出现严重发散，RMSE 达到 920.67 m。该结果表明，仅采用双尺度 Mamba-CNN 特征提取并不足以保证极端城市峡谷场景下的递推稳定性，动态门控

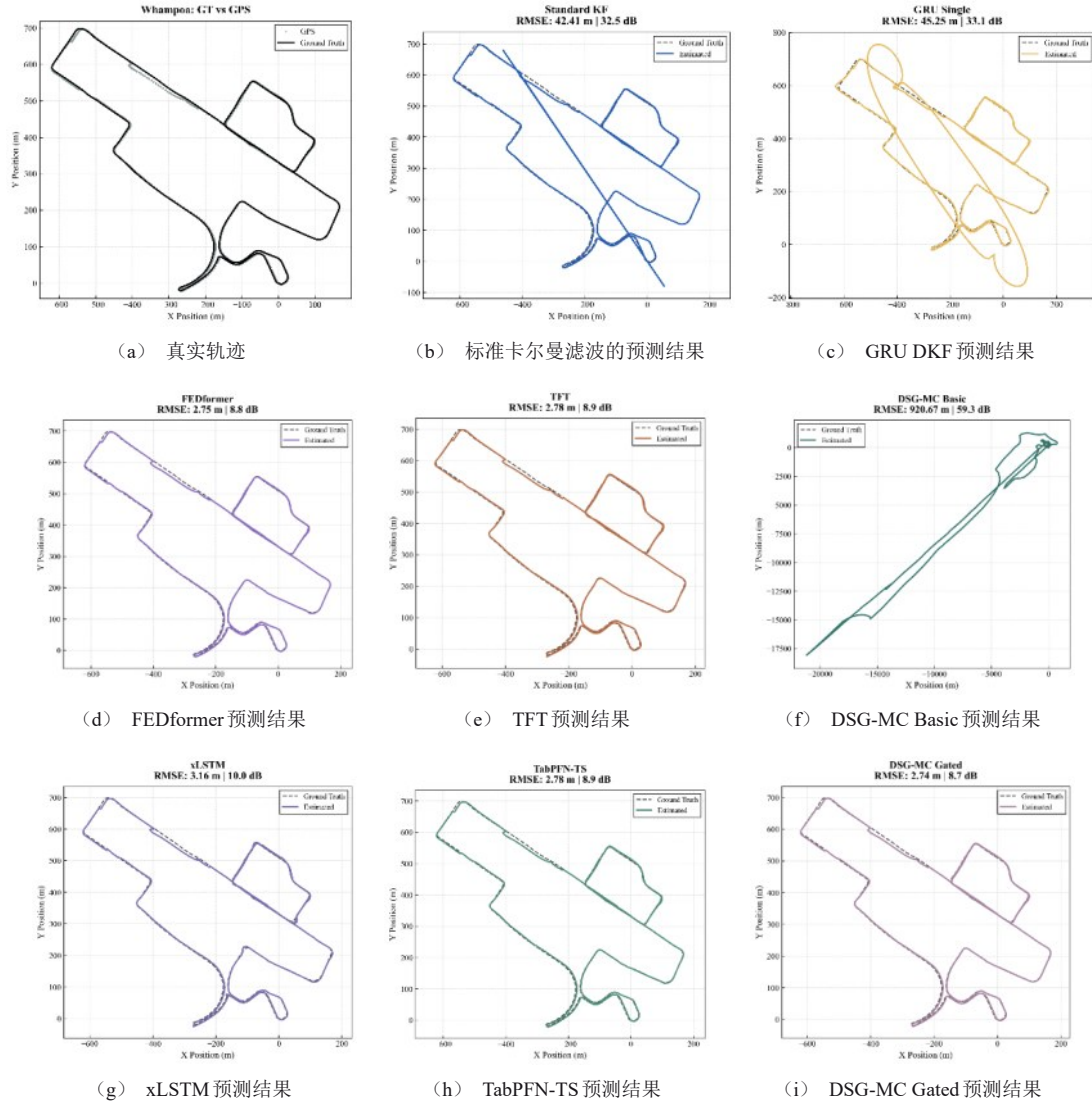


图 13 各模型在 UrbanNav Whampoa 路线下的轨迹估计结果

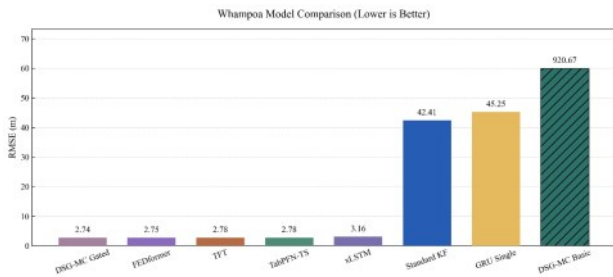


图 14 各个模型在 UrbanNav Whampoa 数据下的实验结果

机制对于抑制异常观测冲击、协调长程与短程特征贡献具有关键作用。综合两条真实路线可以看出, DSG-MC Gated 分别取得 2.78 m 和 2.74 m 的最低 RMSE, 两条路线之间仅相差 0.04 m, 表现出较为稳定的跨路线状态估计性能。

为进一步分析所提模型在训练分布之外的性能退化规律, 本文在保持网络结构、模型参数和训练设置不变的条件下, 开展不同分布偏移条件下的经验学习边界测试。实验从验证集中随机选取 256 条测试序列, 不对网络进行重新训练或目标场景微调, 仅对测试数据中的加速度、转弯率、GNSS 观测噪声及连续 GNSS 拒止时长进行控制变量调整。其中, 加速度、转弯率和 GNSS 噪声的缩放系数均设置为 0.5、1.0、1.5 和 2.0, 连续 GNSS 拒止时长分别设置为 0、1、3、5 和 10s。加速度、转弯率和 GNSS 噪声实验以尺度系数 1.0 作为参考条件, 拒止时长实验则以无 GNSS 拒止条件作为参考。

为统一描述不同压力条件下的性能退化程度, 本文采用绝对位置 RMSE 和相对 RMSE 比值两类指

标进行评价。相对 RMSE 比值定义为

$$r(s) = \frac{\text{RMSE}(s)}{\text{RMSE}(s_{\text{ref}})} \quad (56)$$

其中, s 表示当前测试强度, s_{ref} 表示对应实验的参考条件。为便于定量识别模型性能出现明显退化的位置, 将测试强度增加过程中首次满足 $r(s) \geq 1.5$ 的条件定义为当前测试范围内的经验性能边界。需要指出的是, 该判据主要用于刻画模型在不同分布偏移条件下的性能变化趋势, 并不构成严格的理论稳定性保证。实验结果如图 15 所示。

从运动状态压力测试结果可以看出, 所提模型对一定范围内的加速度和转弯率变化表现出相对平缓的性能退化。当加速度尺度由参考条件的 1.0 倍增加至 1.5 倍和 2.0 倍时, 位置 RMSE 分别由约 0.71 m 增加至 0.78 m 和 0.87 m, 对应相对 RMSE 比值约为 1.10 和 1.23。当转弯率尺度由 1.0 倍增加至 1.5 倍和 2.0 倍时, 位置 RMSE 分别增加至约 0.75

m 和 0.80 m, 对应相对 RMSE 比值约为 1.06 和 1.12。上述两类测试在当前最高 2.0 倍尺度下均未达到 1.5 倍的经验边界阈值, 说明双尺度网络对一定程度的运动强度变化具有较好的适应能力。

与运动状态变化相比, 模型对 GNSS 观测噪声分布偏移更加敏感。当 GNSS 噪声尺度为 0.5 倍时, 位置 RMSE 约为 0.43 m; 在参考噪声条件下, 位置 RMSE 约为 0.71 m。当 GNSS 噪声尺度增加至 1.5 倍时, 位置 RMSE 上升至约 1.24 m, 相对 RMSE 比值约为 1.75, 已超过设定的经验边界阈值; 当噪声尺度进一步增加至 2.0 倍时, 位置 RMSE 增大至约 5.9

m, 相对 RMSE 比值超过 8。该结果表明, 当 GNSS 量测噪声显著超出训练阶段覆盖的统计范围后, 观测新息的幅值和分布特征发生明显变化, 网络根据历史特征预测的数据驱动增益难以继续保持与分布内条件相同的状态修正性能。

连续 GNSS 拒止时长测试进一步反映了模型在缺少绝对量测条件下的性能边界。相较于无拒止参考条件, 1 s 和 3 s 拒止条件下的位置 RMSE 分别约为 0.75 m 和 0.97 m, 对应相对 RMSE 比值约为 1.05 和 1.36, 表明短时间拒止所造成的误差增长相对有限。当连续拒止时长增加至 5 s 时, 位置 RMSE 上升至约 1.64 m, 相对 RMSE 比值约为 2.29, 超过经验边界阈值; 当拒止时长达至 10 s 时, 位置 RMSE 进一步增加至约 6.2 m, 相对 RMSE 比值约为 8.7。该变化趋势与第 3.3 节中的有限拒止区间误差传播分析一致, 即显式量测门控能够阻断无效 GNSS 新息对状态更新的直接污染, 但不能消除标称状态传播误差和未建模扰动随拒止时间的持续累积。

综合上述实验结果, 在本文设定的测试范围和 $r(s) \geq 1.5$ 判据下, 加速度和转弯率方向的经验性能边界均高于当前测试的 2.0 倍尺度, GNSS 噪声方向的经验边界出现在 1.5 倍尺度附近, 连续 GNSS 拒止方向的经验边界出现在 5 s 附近。这说明所提模型对中等程度的运动状态分布偏移具有相对较好的适应能力, 但对明显增大的 GNSS 观测噪声和长时间绝对量测缺失更为敏感。Mamba 长程分支能够利用历史时序信息延缓误差累积, 但无法突破训练数据覆盖范围和观测信息可用性所决定的基

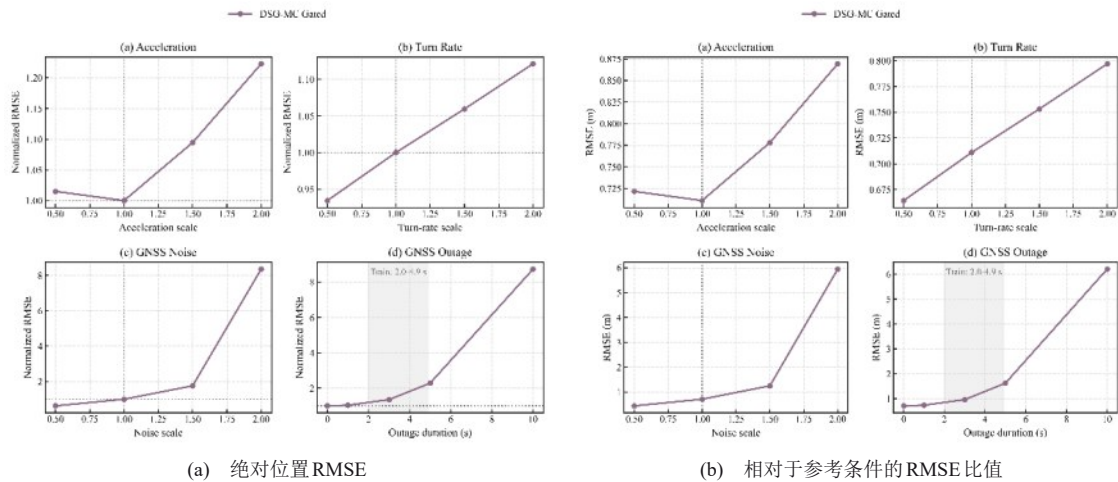


图 15 不同分布偏移条件下 DSG-MC Gated 的经验学习边界测试

本边界。因此,本文关于模型鲁棒性和零样本迁移能力的结论应限定于当前训练分布、运动模式、噪声范围及 GNSS 拒止时长条件,而不能解释为对任意分布外场景的无条件泛化或全局稳定性保证。

在计算效率方面,考虑到自动驾驶和融合定位系统对实时推理延迟的要求,本文在统一测试条件下统计了各模型的推理耗时,结果如表 1 所示。传统 GRU-DKF 采用循环神经网络作为时序骨干,状态信息需要沿时间步顺序递推,推理时延达到 74 ms。统一实现框架下 EKF 的测试时延为 73 ms,该结果反映的是包含状态预测、协方差递推、矩阵求解及程序调度开销在内的工程实现耗时,并不等同于经典 EKF 核心滤波步骤的理论计算复杂度。相比之下,DSG-MC Basic 和 DSG-MC Gated 的推理时延分别为 17 ms 和 15 ms,较 GRU-DKF 分别降低约 77.0% 和 79.7%,其中完整的 DSG-MC Gated 实现了约 4.9 倍的推理加速。这表明,Mamba 的并行扫描机制能够有效缓解传统循环增益网络顺序递推所带来的计算开销。

在新增的时序模型中,TabPFN-TS、TFT、FEDformer 和 xLSTM 的推理时延分别为 9、10、13 和 14 ms。其中,TabPFN-TS 在当前窗口规模和测试设置下取得最低时延,TFT 与 FEDformer 也表现出较高的并行计算效率;xLSTM 虽然仍采用循环记忆结构,但其 14 ms 的推理时延已接近 DSG-MC Gated。需要指出的是,上述模型主要通过窗口化序列直接预测目标状态,而 DSG-MC Gated 还需要依次完成标称状态传播、双尺度特征提取、动态门控融合、增益矩阵生成以及观测新息更新。因此,DSG-MC Gated 并非表 1 中时延最低的模型,但其 15 ms 的推理时延仅比 xLSTM 和 FEDformer 分别增

加 1 ms 和 2 ms,并显著低于 GRU-DKF。同时,结合两条真实城市路线的实验结果,DSG-MC Gated 在保持较低推理时延的基础上,还具备卡尔曼预测—更新递归结构、显式 GNSS 断联门控和观测恢复期异常抑制能力。因此,在当前实验条件下,所提方法在状态估计精度、复杂环境鲁棒性、结构可解释性和计算效率之间表现出较好的综合平衡。

为进一步验证 DSG-MC 中各关键模块对定位精度与断联鲁棒性的实际贡献,本文在保持训练设置和测试数据一致的条件下,对模型中的长程 Mamba 分支、硬测量门控、差分特征清洗、恢复期历史记忆以及 IMU 输入进行逐项消融。考虑到真实城市峡谷场景下误差主要集中在 GNSS 拒止及信号恢复阶段,除整体 RMSE 外,本文进一步统计正常观测区间、断联区间和恢复区间的 RMSE,以更细致地分析不同模块对长序列递推稳定性的影响。相关结果如表 2 所示。

表中给出了 DSG-MC 的核心消融实验结果。可以看出,完整模型 A0 在整体 RMSE 上取得 3.063 m 的定位误差。去除长程 Mamba 分支后,整体 RMSE 上升至 3.239 m,且断联区间 RMSE 由 18.999 m 增加至 21.369 m,说明长程状态空间分支有助于建模惯性序列中的缓变运动趋势,并提升 GNSS 拒止条件下的轨迹连续性。去除硬测量门控后,断联区间 RMSE 显著上升至 82.945 m,表明在 GNSS 缺失阶段若继续执行量测更新,异常新息会严重破坏滤波闭环稳定性。去除差分特征清洗后,整体 RMSE 上升至 6.172 m,说明对观测差分异常值进行清洗能够有效抑制瞬态噪声传播。去除恢复期历史记忆后,整体 RMSE 增加至 52.491 m,验证了恢复期记忆机制对抑制 GNSS 恢复瞬间新息

表 1

各模型推理时延表

模型	时延
EKF	73ms
GRU-DKF	74ms
DSG-MC Basic	17ms
DSG-MC Gated	15ms
TFT	10ms
FEDformer	13ms
xLSTM	14ms
TabPFN-TS	9ms

DSG-MC Basic 表示去除动态门控融合模块后的 DSG-MC DKF 变体;DSG-MC Gated 表示本文提出的完整 DSG-MC DKF 模型。

尖峰和长序列误差放大的重要作用。去除 IMU 输入后, 整体误差和断联区间误差均有所增加, 说明惯性信息对维持断联期间的运动趋势具有辅助作用。综上, 硬测量门控、恢复期记忆、差分特征清洗与长程 Mamba 分支共同构成了 DSG-MC 在复杂城市峡谷场景下保持鲁棒定位的关键机制。

5 结束语

本文提出了 DSG-MC DKF, 一种将 Mamba 结构化状态空间模型与 CNN 结合的新型可微卡尔曼滤波器融合定位算法。不同于传统 GNSS/INS 融合方法将 IMU 作为显式控制输入注入预测方程, 本文通过将控制输入矩阵置零, 取消对固定运动学控制项的依赖, 并利用 Mamba-CNN 网络从 IMU 历史序列、观测新息和断联掩码中隐式学习状态修正规律。实验结果表明, 该策略能够有效缓解动力学模型失配、传感器零偏漂移和 GNSS 长时拒止条件下的状态发散问题, 同时显著降低推理延迟。

一寸证件照

表 2

DSG-MC DKF 消融实验结果

模型变体	Overall	Δ	Normal	Outage	Recovery
Full DSG-MC DKF	3.063	0.000	2.201	18.999	19.074
w/o Long Mamba	3.239	0.176	2.177	21.369	14.345
w/o Hard Gate	9.758	6.695	2.773	82.945	33.921
w/o Feature Cleaning	6.172	3.109	5.804	19.479	3.833
w/o Recovery Memory	52.491	49.428	52.780	19.919	18.686
w/o IMU Input	3.150	0.087	2.170	20.339	11.344

表中数值单位均为 m。 Δ 表示相较于完整模型的 Overall RMSE 增量。

参考文献:

- [1] Kalman R E. A new approach to linear filtering and prediction problems [J]. *Journal of Basic Engineering*, 1960, 82(1): 35-45.
- [2] Julier S J, Uhlmann J K. Unscented filtering and nonlinear estimation [J]. *Proceedings of the IEEE*, 2004, 92(3): 401-422.
- [3] Arasaratnam I, Haykin S. Cubature Kalman filters[J]. *IEEE Transactions on Automatic Control*, 2009, 54(6): 1254-1269.
- [4] 赵方, 吴凡. GNSS/INS 自适应智能组合导航算法[J]. *北京邮电大学学报*, 2022, 45(2): 1-8.
Zhao F, Wu F. Adaptive intelligent GNSS/INS integrated navigation algorithm[J]. *Journal of Beijing University of Posts and Telecommunications*, 2022, 45(2): 1-8.
- [5] 孟秀云, 王语嫣. 一种 SINS/GPS 紧组合导航系统的改进自适应扩展卡尔曼滤波算法[J]. *北京理工大学学报(自然科学版)*, 2018, 38(6): 625-630, 636.
Meng X Y, Wang Y Y. An improved adaptive extended Kalman filtering algorithm for SINS/GPS tightly integrated navigation system[J]. *Transactions of Beijing Institute of Technology*, 2018, 38(6): 625-630, 636.
- [6] 崔冰波, 吉峰, 孙宇, 等. 高斯过程改进的鲁棒容积卡尔曼滤波及其组合导航应用[J]. *电子测量与仪器学报*, 2021, 35(9): 34-40.
Cui B B, Ji F, Sun Y, et al. Gaussian process enhanced robust cubature Kalman filter and application in integrated navigation[J]. *Journal of Electronic Measurement and Instrumentation*, 2021, 35(9): 34-40.
- [7] 沈凯, 刘庭欣, 左思琪, 等. 复杂城市环境下 GNSS/INS 组合导航可观测度分析及鲁棒滤波方法[J]. *仪器仪表学报*, 2020, 41(9): 252-261.
Shen K, Liu T X, Zuo S Q, et al. Observability analysis and robust filtering method for GNSS/INS integrated navigation in complex urban environments[J]. *Chinese Journal of Scientific Instrument*, 2020, 41(9): 252-261.
- [8] 倪少杰, 李诗扬, 谢郁辰, 等. GNSS/INS 超紧组合导航综述[J]. *国防科技大学学报*, 2023, 45(5): 48-59.
Ni S J, Li S Y, Xie Y C, et al. Review of GNSS/INS ultra-tightly integrated navigation[J]. *Journal of National University of Defense Technology*, 2023, 45(5): 48-59.
- [9] 田哲铭, 李旭, 胡悦, 等. 城市峡谷下视觉辅助的 GNSS/INS 多阶段定位方法[J]. *仪器仪表学报*, 2024, 45(4): 217-225.
Tian Z M, Li X, Hu Y, et al. Vision-aided GNSS/INS multi-stage positioning method in urban canyon environments[J]. *Chinese Journal of Scientific Instrument*, 2024, 45(4): 217-225.
- [10] 何康辉, 董朝阳, 王青, 等. 一种考虑 GPS 信号中断的导航滤波算法[J]. *北京航空航天大学学报*, 2019, 45(9): 1874-1881.
He K H, Dong C Y, Wang Q, et al. A navigation filtering algorithm considering GPS signal interruption[J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2019, 45(9): 1874-1881.
- [11] Clark R, Wang S, Wen H, et al. VINet: visual-inertial odometry as a sequence-to-sequence learning problem[C]//*Proceedings of the AAAI Conference on Artificial Intelligence*. Palo Alto: AAAI Press, 2017: 3995-4001.
- [12] Chen C, Rosa S, Miao Y, et al. Selective sensor fusion for neural visual-inertial odometry[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Piscataway: IEEE Press, 2019: 10542-10551.
- [13] Chen C, Lu C X, Markham A, et al. IONet: learning to cure the curse of drift in inertial odometry[C]//*Proceedings of the AAAI Conference on Artificial Intelligence*. Palo Alto: AAAI Press, 2018, 32(1): 1467-1474.
- [14] Herath S, Yan H, Furukawa Y. RoNIN: robust neural inertial navigation in the wild: benchmark, evaluations, and new methods[C]//*Proceedings of the 2020 IEEE International Conference on Robotics and Automation*. Piscataway: IEEE Press, 2020: 3146-3152.
- [15] Liu W, Caruso D, Ilg E, et al. TLIO: tight learned inertial odometry[J]. *IEEE Robotics and Automation Letters*, 2020, 5(4): 5653-5660.
- [16] Cioffi G, Bauersfeld L, Kaufmann E, et al. Learned inertial odometry for autonomous drone racing[J]. *IEEE Robotics and Automation Letters*, 2023, 8(5): 2684-2691.
- [17] 司书斌, 赵大伟, 徐婉莹, 等. 视觉-惯性导航定位技术研究进展[J]. *中国图象图形学报*, 2021, 26(6): 1470-1482.
Si S B, Zhao D W, Xu W Y, et al. Research progress of visual-inertial navigation and positioning technology[J]. *Journal of Image and Graphics*, 2021, 26(6): 1470-1482.
- [18] 施俊屹, 查富生, 孙立宁, 等. 移动机器人视觉惯性 SLAM 研究进展[J]. *机器人*, 2020, 42(6): 734-748.
Shi J Y, Zha F S, Sun L N, et al. Research progress of visual-inertial SLAM for mobile robots[J]. *Robot*, 2020, 42(6): 734-748.
- [19] 王金科, 左星星, 赵祥瑞, 等. 多源融合 SLAM 的现状与挑战[J]. *中国图象图形学报*, 2022, 27(2): 368-389.
Wang J K, Zuo X X, Zhao X R, et al. Status and challenges of multi-source fusion SLAM[J]. *Journal of Image and Graphics*, 2022, 27(2): 368-389.
- [20] LIM B, ARIK S Ö, LOEFF N, et al. Temporal Fusion Transformers for interpretable multi-horizon time series forecasting[J]. *International Journal of Forecasting*, 2021, 37(4): 1748-1764.
- [21] ZHOU T, MA Z, WEN Q, et al. FEDformer: Frequency enhanced decomposed transformer for long-term series forecasting[C]//*Proceedings of the 39th International Conference on Machine Learning*. PMLR, 2022: 27268-27286.
- [22] WANG S, WU H, SHI X, et al. TimeMixer: Decomposable multiscale mixing for time series forecasting[C]//*International Conference on Learning Representations*. 2024.
- [23] AHAMED M A, CHENG Q S. TimeMachine: A time series is worth 4 Mambas for long-term forecasting[C]//*Proceedings of the 27th European Conference on Artificial Intelligence*. 2024: 1688-1695.
- [24] WANG Z, KONG F, FENG S, et al. Is Mamba effective for time series forecasting[J]. *Neurocomputing*, 2025, 619: 129178.
- [25] BRIGATO L, MORAND R, STRØMMEN K J, et al. There are no champions in supervised long-term time series forecasting[J]. *Transactions on Machine Learning Research*, 2026.
- [26] Haarnoja T, Ajay A, Levine S, et al. Backprop KF: learning discriminative deterministic state estimators[C]//*Advances in Neural Information Processing Systems* 29. Red Hook: Curran Associates, 2016: 4376-4384.
- [27] Revach G, Shlezinger N, Ni X, et al. KalmanNet: neural network aided Kalman filtering for partially known dynamics[J]. *IEEE Transactions on Signal Processing*, 2022, 70: 1532-1547.
- [28] Bengio Y, Simard P, Frasconi P. Learning long-term dependencies with gradient descent is difficult[J]. *IEEE Transactions on Neural Networks*, 1994, 5(2): 157-166.
- [29] Cho K, van Merriënboer B, Gulcehre C, et al. Learning phrase repre-

- sentations using RNN encoder-decoder for statistical machine translation[C]//Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing. Stroudsburg: Association for Computational Linguistics, 2014: 1724-1734.
- [30] Gu A, Goel K, Ré C. Efficiently modeling long sequences with structured state spaces[C]//International Conference on Learning Representations. San Juan: ICLR, 2022: 1-21.
- [31] Gu A, Dao T. Mamba: linear-time sequence modeling with selective state spaces[C]//Proceedings of the Conference on Language Modeling. Philadelphia: COLM, 2024: 1-21.
- [32] Lea C, Flynn M D, Vidal R, et al. Temporal convolutional networks for action segmentation and detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2017: 156-165.
- [33] Lorenz E N. Deterministic nonperiodic flow[J]. Journal of the Atmospheric Sciences, 1963, 20(2): 130-148.
- [34] Carlevaris-Bianco N, Ushani A K, Eustice R M. University of Michigan North Campus long-term vision and lidar dataset[J]. The International Journal of Robotics Research, 2016, 35(9): 1023-1035.
- [35] Tobin J, Fong R, Ray A, et al. Domain randomization for transferring deep neural networks from simulation to the real world[C]//Proceedings of the 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway: IEEE Press, 2017: 23-30.
- [36] BECK M, PÖPPEL K, SPANRING M, et al. xLSTM: Extended long short-term memory[C]//Advances in Neural Information Processing Systems. 2024, 37.
- [37] HOLLMANN N, MÜLLER S, PURUCKER L, et al. Accurate predictions on small data with a tabular foundation model[J]. Nature, 2025, 637(8045): 319-326. DOI: 10.1038/s41586-024-08328-6.
- [38] Hsu L T, Huang F, Ng H F, et al. Hong Kong UrbanNav: an open-source multisensory dataset for benchmarking urban navigation algorithms[J]. Navigation, 2023, 70(4): navi.602.

姓名（出生年— ），性别，籍贯（如陕西西安人），博士，单位职称（如西安邮电大学教授），主要研究方向为公钥加密、云安全和无线网络安全。