arXiv:2303.01291v1 [cs.RO] 2 Mar 2023

通过视觉惯性融合实现稳健、高精度的GNSS载波相位定位

Erqun Dong[1, 2*], Sheroze Sheriffdeen[2], Shichao Yang[2], Jing Dong[2], Renzo De Nardi[2], Carl Ren[2], Xiao-Wen Chang[1], Xue Liu[1] 和 Zijian Wang[2]

摘要——稳健、高精度的全球定位是众多户外机器人应用的基础。传统融合方法采用精度较低的基于伪距的GNSS测量值（$> 5m$ 误差），只能实现相对于全球地心地固（ECEF）坐标系的粗略配准。本文利用高精度GNSS载波相位定位，并在扩展卡尔曼滤波器（EKF）框架下，借助局部视觉惯性里程计（VIO）跟踪，以更好地求解GNSS载波相位涉及的整周模糊度。我们还提出了一种算法，用于精确标定GNSS天线到IMU的外参，从而将VIO准确对齐至ECEF坐标系。通过在严重遮挡的城市峡谷中进行真实环境硬件实验，我们的系统展现出稳健的全球定位能力；在整周模糊度固定率和定位均方根误差精度方面，显著优于当前最先进的RTKLIB。

I. 引言

本文借助视觉惯性里程计，解决复杂城市峡谷环境下基于GNSS载波相位的全球定位鲁棒性问题。

鲁棒的厘米级全球定位是移动机器人、自动驾驶汽车和增强现实等诸多户外应用的基石。目前，许多定位方案依赖视觉惯性里程计（VIO） [1], [2], [3]，因为它能够准确地进行短距离跟踪。由于缺少预先构建的地图（而地图的构建和维护成本高昂），VIO在需要精确全球定位的大规模应用中用途有限。受近期一些将GNSS伪距定位（GNSS-PR）融合到VIO中以减轻VIO漂移的工作的启发 [4], [5], [6], [7], [8]，我们将VIO融合到GNSS载波相位定位（GNSS-CP）中，旨在为户外应用实现低成本、高精度的全球定位。

GNSS-CP已展现出在开阔天空下实现无漂移、精度达到 2–5cm 的厘米级定位能力。相比之下，精度为 >5 m 的 GNSS-PR 在厘米级定位方面潜力有限。示例见图 1。然而，在密集城区，GNSS-CP更容易受到视线遮挡和多路径干扰的影响，因而更加脆弱。第 V-D 节表明，在城市峡谷中，GNSS-CP的精度可能从厘米级降至数十米。因此，为实现稳健的 GNSS

*本工作是在 Meta 实习期间完成的。[1]麦吉尔大学计算机科学学院，加拿大魁北克省蒙特利尔，erqun.dong@mail.mcgill.ca，{chang, xueliu}@cs.mcgill.ca [2]Meta，Reality Labs Research，美国华盛顿州雷德蒙德{sheroze, shichaoy, jingdong, renzo, carlren, zjwang}@fb.com



图 1：全球导航卫星系统伪距轨迹（红色）、GNSS载波相位轨迹（绿色）以及在前 50 米内与 GNSS载波相位对齐的视觉惯性里程计轨迹（蓝色）示意图。浅绿色数字和虚线表示轨迹四个主要部分的行进方向。可以观察到，伪距轨迹比另外两条轨迹噪声明显更大，定位误差最高可达 10 米。这表明，与全球导航卫星系统伪距相比，GNSS载波相位有望实现更精确的全球定位。

载波相位全球定位，应对复杂城市峡谷环境，本文重点解决以下问题：1）利用 VIO 的局部跟踪和不确定性估计，更好地调整并约束整数最小二乘的搜索空间；2）从工程角度出发，剔除城市环境中的异常卫星测量值，从而降低 GNSS 载波相位解算受噪声多路径信号影响的可能性。

为此，我们采用所提出的标定算法，将视觉惯性里程计轨迹仔细配准到 ECEF 坐标系，以求解惯性测量单元与GNSS天线中心之间的平移偏移量。我们在扩展卡尔曼滤波框架中松耦合视觉惯性里程计与 GNSS-CP，其中视觉惯性里程计作为运动模型，用于传播先验状态估计值及协方差；GNSS-CP 则作为测量更新，以获得后验浮点解。我们设计了严格的异常值剔除机制，以应对城市峡谷等具有挑战性的环境中的信号遮挡和多路径反射。总而言之，通过良好的浮点解（其中整周模糊度被视为实数）作为初始值，可以进行整周模糊度解算，进而求得固定解（其中整周模糊度被固定为整数），并提高成功率。

我们在城市峡谷中使用 Aria眼镜 [9] 开展硬件实验，并为其刚性安装了一款低成本、多频段 GNSS天线。结果表明，我们的标定

可实现厘米级对齐误差；此外，我们方法的总体定位性能和整周模糊度固定率均优于名为 RTKLIB [10] 的先进 GNSS-CP 方法。值得一提的是，从建筑物高度和密度来看，我们实验中的城市峡谷（图 4、5）比近期论文 [11], [12], [13], [7] 中的实验环境更为重度。

II. 相关工作

自视觉与视觉惯性定位和建图取得成功以来，视觉惯性信息与全球导航卫星系统的融合便已得到研究 [1], [14], [2]。一些工作 [15], [6], [4], [5], [12], [8]通过融合视觉惯性信息与全球导航卫星系统伪距定位，在局部ENU坐标系中进行定位。另一些工作则利用全球导航卫星系统伪距测量，将局部里程计坐标系对齐到全球ECEF坐标系 [16], [7]。这些方法将伪距定位结果或伪距测量建模为高斯噪声。然而，由于全球导航卫星系统伪距测量/定位噪声较大，即使在开阔天空下，视觉惯性里程计与全球导航卫星系统的对齐仍会产生数米范围内的较大误差（参见 [7] 中的表 II）。值得注意的是，大多数融合伪距 GNSS 测量与视觉惯性传感（如 [5], [6], [8]）的工作，都以普通实时动态定位或 RTKLIB 作为真值进行结果对比；而本文取得了优于其他工作所采用的实时动态定位真值的性能。

已有研究尝试结合 GNSS载波相位、相机和惯性测量单元实现全球定位。Li等人 [11]和 Yoder等人 [13]将视觉惯性信息融合为先验估计，以辅助整周模糊度解算。Shepard等人 [17]采用光束法平差方法进行估计。这些论文均假设惯性测量单元与天线外参标定参数是预先已知的常量，却未对其取值进行表征。相比之下，我们提出了一种轻量级优化方法来估计这些参数，并通过实验对其进行验证。此外，这三项工作都缺少像我们这样的、在极具挑战性的城市峡谷中开展的真实场景实验。此外，[13] 的局限在于要求在设备上安装两根 GNSS天线。

已有研究致力于在天空可视度不佳的情况下辅助载波相位定位。Takasu等人 [18] 采用惯性测量单元进行航位推算，以接管导航。利用激光雷达 [19] 构建的建筑物高度来减少卫星信号离群值。研究人员还采用基于阈值的离群值检测方法来缓解具有挑战性的城市峡谷效应 [11], [12], [13], [7], ，但其真实场景实验环境并不像我们的环境那样具有挑战性。在本文中，我们将 VIO 的增量运动视为强先验，并在测量更新和整周模糊度解算中剔除离群值。

III. 问题表述

假设我们的 GNSS 模块和参考站接收到 g 组信号。每组信号都属于同一星座且频率相同，并且假设第 j 组有 m_j 颗卫星。那么 GNSS测量在某一历元为 $\boldsymbol{y} = [\boldsymbol{\phi}_1^\top, \ldots, \boldsymbol{\phi}_g^\top, \boldsymbol{\rho}_1^\top, \ldots, \boldsymbol{\rho}_g^\top]^\top$，其中$\boldsymbol{\phi}_j$是第 j组的载波相位双差测量 [20]，$\boldsymbol{\rho}_j$ 是第 j组的伪距双差测量。选取一颗卫星作为双差的参考卫星后，$\boldsymbol{\phi}_j$和$\boldsymbol{\rho}_j$ 的维度均为 $(m_j - 1)$。

视觉惯性输入由VIO处理，提供以惯性测量单元为中心、位于局部里程计坐标系 {O} 中的设备位姿 {I}，记为 $\boldsymbol{T}_{\mathrm{OI}} \in$ 三维特殊欧几里得群，其中 SE(3) 表示三维特殊欧几里得群。VIO还会输出相应的不确定性 [1]。随后，我们在全球导航卫星系统载波相位测量值的时间戳处对VIO轨迹进行插值，以获得时间插值后的VIO位姿。我们采用松耦合融合方法，将VIO的输出单向输入GNSS-CP算法。我们假设相机和惯性测量单元的内参与外参均已标定，而GNSS天线与惯性测量单元之间的外参（记为 $\boldsymbol{t}_{\mathrm{IA}} \in \mathbb{R}^3$）需要估计。

我们假设，在融合算法启动之前，载波相位定位已初步收敛（不一定达到固定解），这使我们能够将VIO局部里程计坐标系与ECEF坐标系初步对齐（如图4b和图5b所示），从而为算法提供热启动。

待估计的状态表示为 $\boldsymbol{x} = [{}^{\mathrm{E}}\boldsymbol{p}_{\mathrm{u}}^\top \ \boldsymbol{d}_1^\top, \ \ldots, \ \boldsymbol{d}_g^\top]^\top$，其中 ${}^{\mathrm{E}}\boldsymbol{p}_{\mathrm{u}}$ 表示用户在 ECEF 坐标系 {E}中的当前位置，而 $\boldsymbol{d}_j = [d_j^{(1,2)}, d_j^{(1,3)}, \ldots, d^{(1,m_j)}{}_j \ ]^\top$ 表示组 j中卫星的双差整周模糊度。我们的融合系统基于名为 RTKLIB [10]的基准扩展卡尔曼滤波器构建。基准扩展卡尔曼滤波器的预测步骤采用匀速模型；而我们使用视觉惯性里程计进行更精确的运动学建模，详见第 IV-B 节。对于测量更新步骤，测量方程采用载波相位双差测量和伪距双差测量。具体而言，对于每个组 j，

$$\boldsymbol{\phi}_j = \boldsymbol{h}_{\phi,j}(\boldsymbol{x}) + \boldsymbol{\epsilon}_{\phi,j}, \quad \boldsymbol{\rho}_j = \boldsymbol{h}_{\rho,j}(\boldsymbol{x}) + \boldsymbol{\epsilon}_{\rho,j},$$

其中，$\boldsymbol{\epsilon}_{\phi,\ j}$ 是载波相位双差测量的高斯噪声，$\boldsymbol{\epsilon}_{\rho,\ j}$ 是伪距双差测量的高斯噪声（关于$\boldsymbol{\epsilon}_{\phi,\ j}$和 $\boldsymbol{\epsilon}_{\rho,\ j}$的公式，读者可参阅第 7.5 节）。$\boldsymbol{h}_{\phi,\ j}$中的各个测量函数（$\boldsymbol{x}$）和$\boldsymbol{h}_{\rho,\ j}$（$\boldsymbol{x}$）为

$$h_{\phi,j}^{(1,s)}(\boldsymbol{x}) = (||{}^{\mathrm{E}}\boldsymbol{t}_{\mathrm{u}}^{(1)}||_2 - ||{}^{\mathrm{E}}\boldsymbol{t}_{\mathrm{r}}^{(1)}||_2) - (||{}^{\mathrm{E}}\boldsymbol{t}_{\mathrm{u}}^{(s)}||_2 - ||{}^{\mathrm{E}}\boldsymbol{t}_{\mathrm{r}}^{(s)}||_2) + \lambda_j d_j^{(1,s)},$$

$$h_{\rho,j}^{(1,s)}(\boldsymbol{x}) = (||{}^{\mathrm{E}}\boldsymbol{t}_{\mathrm{u}}^{(1)}||_2 - ||{}^{\mathrm{E}}\boldsymbol{t}_{\mathrm{r}}^{(1)}||_2) - (||{}^{\mathrm{E}}\boldsymbol{t}_{\mathrm{u}}^{(s)}||_2 - ||{}^{\mathrm{E}}\boldsymbol{t}_{\mathrm{r}}^{(s)}||_2),$$

其中，下标 u 和 r 分别表示用户和参考站，上标 1 和 s 表示卫星标识符。不失一般性，我们假设参考卫星的卫星标识符为 1；λ_j 表示组 j 信号的波长；$||{}^{\mathrm{E}}\boldsymbol{t}^{(1)}{}_{\mathrm{u}}\ ||_2$ 表示用户 u 与满足 $||{}^{\mathrm{E}}\boldsymbol{t}^{(1)}{}_{\mathrm{u}}\ ||_2 = ||{}^{\mathrm{E}}\boldsymbol{p}_{\mathrm{u}} - {}^{\mathrm{E}}\boldsymbol{p}^{(1)}||_2$ 的卫星 1 之间的距离；$||{}^{\mathrm{E}}\boldsymbol{t}^{(1)}{}_{\mathrm{r}}\ ||_2$、$||{}^{\mathrm{E}}\boldsymbol{t}^{(s)}{}_{\mathrm{u}}\ ||_2$ 和 $||{}^{\mathrm{E}}\boldsymbol{t}^{(s)}{}_{\mathrm{r}}\ ||_2$ 的含义类似。

扩展卡尔曼滤波测量更新将 $\boldsymbol{x}$ 的所有分量（包括位置变量和整周模糊度）都视为实数，以求得浮点解 [20]。基于浮点解，我们随后使用 LAMBDA算法 [21], [22]

来求解固定解，其中一个核心步骤是求解（搜索）整数最小二乘问题

$$\boldsymbol{D} = \underset{\boldsymbol{D}\in\mathbb{Z}^N}{\operatorname{argmin}} \ (\boldsymbol{D}-\hat{\boldsymbol{D}})^\top \boldsymbol{W}_{\hat{\boldsymbol{D}}}^{-1}(\boldsymbol{D}-\hat{\boldsymbol{D}}), \tag{1}$$

其中，$\boldsymbol{D} = [\boldsymbol{d}_1^\top,\ldots,\boldsymbol{d}_g^\top]^\top$ 表示所有组中的全部双差ˆ整周模糊度，$\boldsymbol{D}$ 表示其浮点解，N表示其维度，$\boldsymbol{W}^{-1}_{\hat{\boldsymbol{D}}}$ 表示 $\boldsymbol{D}$的ˆ协方差的逆矩阵。经过整周模糊度搜索并通过比值检验 [23], 验证后，我们可以得到<E> $\boldsymbol{p}_{\mathrm{u}}$ 的固定解，其精度高于浮点解 [21]。在这个整数最小二乘问题中，浮点解及其协方差构成了正确整数的椭球搜索空间，因此它们的质量会影响整周模糊度搜索的成功率，进而影响固定解率。

简而言之，本文利用每个历元的GNSS测量值 $\boldsymbol{y}$ 和视觉惯性里程计位姿 $\boldsymbol{T}_{OI}$，在每个历元计算状态 $\boldsymbol{x}$ 。

IV. 方法

我们通过将视觉惯性里程计融合到GNSS载波相位中，提高式（1）中浮点解和协方差的质量，以实现鲁棒定位和更好的整周模糊度解算。在先验预测和协方差传播方面，对齐良好的视觉惯性里程计都比RTKLIB的匀速模型精确得多。此外，视觉惯性里程计跟踪的精度也能有效排除异常卫星，从而仅使用视距全球卫星导航系统信号进行定位。本节介绍这些方法，并在第 V-C 节和第 V-D 节展示系统性能。

A. 通过外参标定对齐视觉惯性里程计与全球导航卫星系统

我们通过最小二乘优化对齐VIO轨迹与全球卫星导航系统轨迹，从而获得从里程计坐标系{O} 到ECEF坐标系 {E}, $\boldsymbol{T}_{\mathrm{EO}}$的变换。为实现精确对齐，还应考虑外参 $\boldsymbol{t}_{\mathrm{IA}}$，它表示惯性测量单元中心与GNSS天线中心之间的平移。给定由一系列用户三维位置组成的全球导航卫星系统轨迹 ${}^{\mathrm{E}}\boldsymbol{p}_{\mathrm{u}_i}$，$i=1,\ldots,k$，以及与之对应、经时间插值得到的VIO轨迹 $\boldsymbol{T}_{\mathrm{OI}_i}$，$i=1,\ldots,k$，我们对变换 $\boldsymbol{T}_{\mathrm{EO}}$ 和外参 $\boldsymbol{t}_{\mathrm{IA}}$ 进行优化，具体如下

$$\min_{\substack{\boldsymbol{T}_{\mathrm{EO}}\in\mathrm{SE}(3),\\ \boldsymbol{t}_{\mathrm{IA}}\in\mathbb{R}^3}} \ \sum_{i=1}^{k}\left\| \boldsymbol{T}_{\mathrm{EO}}\cdot\boldsymbol{T}_{\mathrm{OI}_i}\cdot\begin{bmatrix}\boldsymbol{t}_{\mathrm{IA}}\\ 1\end{bmatrix} - {}^{\mathrm{E}}\boldsymbol{p}_{\mathrm{u}_i}\right\|^2. \tag{2}$$

为了使对齐 $\boldsymbol{T}_{\mathrm{EO}}$ 在所有自由度上都得到充分约束，轨迹在平移和旋转分量上需要具有持续激励 [24] 。为此，我们使用主成分分析 [25] 判断第一特征值与所有特征值之和的比值（FER）是否小于某个阈值。

我们采用两遍过程求解优化问题式 (2)。首先，将 $\boldsymbol{T}_{\mathrm{EO}}$ 初始化为单位值，将 $\boldsymbol{t}_{\mathrm{IA}}$ 固定为 0，并对式 (2)进行优化，以粗略估计 $\boldsymbol{T}_{\mathrm{EO}}$。其次，我们将 $\boldsymbol{T}_{\mathrm{EO}}$ 初始化为第一遍得到的粗略估计，并再次对式 (2)进行优化，联合求解 $\boldsymbol{T}_{\mathrm{EO}}$ 和 $\boldsymbol{t}_{\mathrm{IA}}$ 。两遍过程是必要的，可确保非线性优化以良好的初始值启动，从而避免陷入不良局部最小值。V-B节展示了我们提出的标定方法在提升对齐精度方面的效果。

考虑到视觉惯性里程计（VIO）会发生漂移，我们需要频繁地将 VIO 重新对齐到 ECEF 坐标系。后续所有对齐都会在固定 $\boldsymbol{t}_{\mathrm{IA}}$ 的情况下优化公式 2 中的 $\boldsymbol{T}_{\mathrm{EO}}$ 。在真实的城市峡谷环境中，GNSS 定位可能会严重退化，影响轨迹对齐的准确性。因此，我们采用 0.1 m 的均方根误差（RMSE）阈值和 97% 的 FER 阈值，来决定是否在新的历元进行轨迹对齐。重新对齐的频率标注在图 4b 和图 5b 中。

B. 融合算法

将 VIO 配准到 ECEF 坐标系后，我们以松耦合方式融合 VIO 与 GNSS。我们利用 VIO 跟踪 ECEF 坐标系中的增量运动，将其作为扩展卡尔曼滤波器（EKF）的先验预测并传播协方差。随后，我们按照 RTKLIB [10] 中的流程，使用 MLAMBDA算法 [22] 求解整周模糊度。本节将介绍系统的主要设计。整体融合流程的伪代码总结于算法1中。

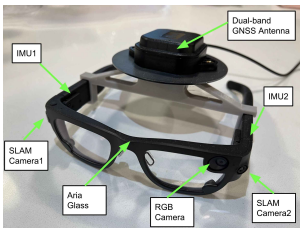
算法1 每个历元的融合算法

```
在历元 i+1 采集新历元数据
ˆ2: T_EO ←AlignNewSegmentationOrReuseOldAlignment()
3: ○ t_{I_i I_{i+1}} ←视觉惯性里程计（新历元数据）
4: E p_{u_{i+1|i}} ←公式（3）
5: P_{i+1|i} ← 公式（6）
6: ȳ ←异常卫星剔除（新历元数据）
7: 浮点解←测量更新(ȳ, x)
8: 如果浮点解验证通过则
9:     固定解←MLAMBDA(ȳ)
10:    如果固定解验证通过则
11:        最终解←固定解和协方差
12:    else
13:        ȳ′ ←进一步剔除产生较大残差的卫星
       整数最小二乘中的值
14:        固定解← MLAMBDA(ȳ′)
15:        如果 固定解验证通过则
16:            最终解←固定解和协方差
17:        else
18:            最终解←浮点解和协方差
19:        结束如果
20:    结束 if
21: 否则
22:    最终解←先验 E p_{u_{i+1|i}}  P_{i+1|i}
23: 结束 if
24: 返回最终解
```

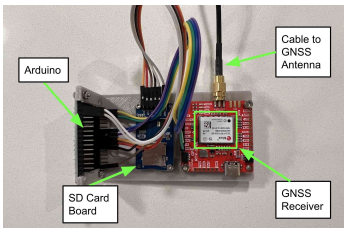
*1) EKF*先验预测： 假设历元 i 的状态已估计，并且我们需要估计历元$i+1$的状态。由于视觉惯性里程计的频率高于全球导航卫星系统，我们在视觉惯性里程计的滑动窗口中寻找采集时间最接近全球导航卫星系统历元 i 和 $i+1$的两个位姿，分别记为 $\boldsymbol{T}_{\mathrm{OI}_i}$和 $\boldsymbol{T}_{\mathrm{OI}_{i+1}}$。随后，可根据前一历元的位置估计${}^{\mathrm{E}}\boldsymbol{p}_{\mathrm{u}_{i|i}}$以及视觉惯性里程计在里程计坐标系中的增量平移${}^{\mathrm{O}}\boldsymbol{t}_{\mathrm{I}_i\mathrm{I}_{i+1}}$，预测先验位置估计${}^{\mathrm{E}}\boldsymbol{p}_{\mathrm{u}_{i+1|i}}$，其表达式为

$${}^{\mathrm{E}}\boldsymbol{p}_{\mathrm{u}_{i+1|i}} = {}^{\mathrm{E}}\boldsymbol{p}_{\mathrm{u}_{i|i}} + \boldsymbol{R}_{\mathrm{EO}}\cdot{}^{\mathrm{O}}\boldsymbol{t}_{\mathrm{I}_i\mathrm{I}_{i+1}}, \tag{3}$$

其中 $\boldsymbol{R}_{\mathrm{EO}}$ 是 $\boldsymbol{T}_{\mathrm{EO}}$的旋转部分。



(a) 配有外置双频GNSS天线的 Aria眼镜。



(b) GNSS接收机板。

图 2：硬件实验中使用的设备。

2）协方差传播：设 ${}^{\mathrm{E}}\boldsymbol{p}_{\mathrm{u}_{i+1}}$ 的先验协方差为 $\boldsymbol{P}_{i+1|i}$，${}^{\mathrm{E}}\boldsymbol{p}_{\mathrm{u}_i}$ 的后验协方差为 $\boldsymbol{P}_{i|i}$。我们结合视觉惯性里程计的协方差，从 $\boldsymbol{P}_{i|i}$ 传播 $\boldsymbol{P}_{i+1|i}$。我们根据视觉惯性里程计维护的联合高斯分布的协方差，得到 $\begin{bmatrix}\boldsymbol{\xi}_{\mathrm{OI}_i}\\ \boldsymbol{\xi}_{\mathrm{OI}_{i+1}}\end{bmatrix}$ 的 12×12 协方差矩阵 $\mathrm{cov}(\begin{bmatrix}\boldsymbol{\xi}_{\mathrm{OI}_i}\\ \boldsymbol{\xi}_{\mathrm{OI}_{i+1}}\end{bmatrix})$，其中 $\boldsymbol{\xi}_{\mathrm{OI}_i}$ 和 $\boldsymbol{\xi}_{\mathrm{OI}_{i+1}}$ 分别是 $\boldsymbol{T}_{\mathrm{OI}_i}$ 和 $\boldsymbol{T}_{\mathrm{OI}_{i+1}}$对应的李代数元素 [26] 。然后，里程计坐标系中增量平移的协方差为

$$\mathrm{cov}({}^{\mathrm{O}}\boldsymbol{t}_{\mathrm{I}_i\mathrm{I}_{i+1}}) = [\boldsymbol{J}_{\mathrm{I}_i}\boldsymbol{J}_{\mathrm{I}_{i+1}}]\,\mathrm{cov}(\begin{bmatrix}\boldsymbol{\xi}_{\mathrm{OI}_i}\\ \boldsymbol{\xi}_{\mathrm{OI}_{i+1}}\end{bmatrix})\begin{bmatrix}\boldsymbol{J}_{\mathrm{I}_i}^{\top}\\ \boldsymbol{J}_{\mathrm{I}_{i+1}}^{\top}\end{bmatrix}, \quad (4)$$

其中，$\boldsymbol{J}_{\mathrm{I}_i}$ 和 $\boldsymbol{J}_{\mathrm{I}_{i+1}}$ 分别是${}_{\mathrm{O}}\boldsymbol{t}_{\mathrm{I}_i\mathrm{I}_{i+1}}$关于 $\boldsymbol{T}_{\mathrm{OI}_i}$ 和 $\boldsymbol{T}_{\mathrm{OI}_{i+1}}$的李代数元素的偏导数。在式 (4) 中，$\mathrm{cov}({}^{\mathrm{O}}\boldsymbol{t}_{\mathrm{I}_i\mathrm{I}_{i+1}})$ 是 3×3，而 $\boldsymbol{J}_{\mathrm{I}i}$和 $\boldsymbol{J}_{\mathrm{I}_{i+1}}$均为 3×6。随后，在地心地固坐标系中，

$$\mathrm{cov}({}^{\mathrm{E}}\boldsymbol{t}_{\mathrm{I}_i\mathrm{I}_{i+1}}) = \boldsymbol{R}_{\mathrm{EO}}\cdot\mathrm{cov}({}^{\mathrm{O}}\boldsymbol{t}_{\mathrm{I}_i\mathrm{I}_{i+1}})\cdot\boldsymbol{R}_{\mathrm{EO}}^{\top}. \quad (5)$$

所提出的扩展卡尔曼滤波预测步骤中 GNSS先验位置的协方差传播的最终形式为

$$\boldsymbol{P}_{i+1|i} = \boldsymbol{P}_{i|i} + \mathrm{cov}({}^{\mathrm{E}}\boldsymbol{t}_{\mathrm{I}_i\mathrm{I}_{i+1}}). \quad (6)$$

3) 异常卫星剔除：实际上，由于GNSS信号在城市峡谷中会受到重度遮挡和反射，沿非视距（NLoS）路径传播的信号不应参与定位。因此，我们需要检测并剔除这类异常卫星测量值。我们会在测量更新步骤和整周模糊度解算步骤中进行异常值剔除。

在测量更新步骤中，由于视觉惯性里程计可提供准确的扩展卡尔曼滤波预测，我们将视觉惯性里程计增量运动的 5 倍用作阈值，以筛选卫星测量值的新息。剩余测量值 $\bar{\boldsymbol{y}}$ 是$\boldsymbol{y}$的子集。此后，$\bar{\boldsymbol{y}}$ 中的某些元素仍可能不可靠，因为部分非视距信号通过了阈值筛选。随后在整周模糊度解算步骤中，如果通过 MLAMBDA [22] 算法解算出的整周模糊度未能通过比值检验 [23], ，系统会进一步剔除在整数最小二乘问题中产生最大残差的卫星，然后重新求解。这样可以排除更多虚假离群值。

V. 实验

A. 硬件实现

我们在真实且具有挑战性的城市峡谷环境（美国华盛顿州贝尔维尤）中进行了硬件实验，使用的是 Aria眼镜 [9]，其配备了一系列机器感知传感器，包括两台全局快门单色SLAM相机、两个六自由度惯性测量单元，以及内置的GPS伪距模块。我们使用金属接地板将一副外接多频GNSS载波相位天线刚性固定在 Aria眼镜上，如图 2a 所示。GNSS载波相位测量值由 u-BLox ZED-F9P GNSS接收机接收，原始数据则由 Arduino 进一步记录到 SD卡中，以供离线处理[1]。实验中使用了三个星座：GPS、GLONASS 和伽利略。我们使用相同的天线和接收机架设了自己的参考站，距用户接收机约 20 公里。在北美的许多城市，公共基站与其附近城市环境之间的基线长度通常如此。两台SLAM相机以 10 Hz 的频率运行，分辨率为 640×480；两个惯性测量单元的运行频率分别为 1000 Hz 和 800 Hz。内置 GPS伪距模块以 1 Hz 运行，外接 GNSS载波相位模块则以 5 Hz 运行。

为了实现GNSS载波相位与VIO之间的同步，由于Aria眼镜的内置GPS和外部GNSS载波相位天线都会记录GPS 时间，我们可以直接使用Aria眼镜的内置GPS作为桥梁，同步VIO（Aria眼镜上的系统时间）与外部GNSS载波相位信号。

B. 对齐精度标定分析

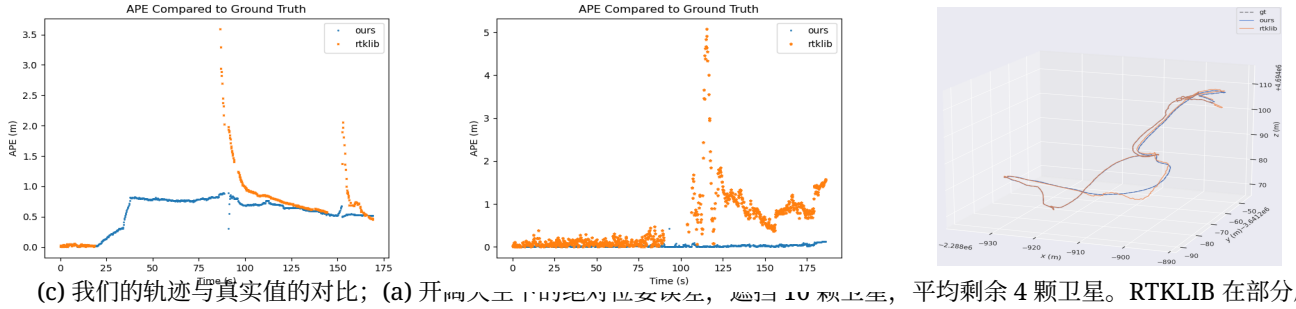
表I展示了在无遮挡天空可视度条件下，针对三条轨迹进行的对齐实验结果，包括使用和不使用所提出的标定方法两种情况；三条轨迹的长度分别为100m、50 m和30m。可以观察到，标定方法将对齐误差从分米级显著降低至厘米级，在轨迹3中最大降低了12cm，从而大幅提升了对齐精度。另一个结论是，采用标定后的对齐误差量级与载波相位定位本身相同，因此VIO与载波相位定位的精度相当，适合进行互补传感器融合。

	轨迹 1	轨迹 2	轨迹 3
对齐长度（米）	100	50	30
未标定时的均方根误差（米）	0.11	0.11	0.16
标定后的均方根误差（米）	**0.071**	**0.028**	**0.037**

表 I：有标定和无标定时的对齐误差

此外，我们通过一项开阔天空下的实地实验表明，GNSS 伪距测量噪声过大，无法为视觉惯性里程计提供可靠的、以 ECEF 为参考的厘米级配准。如图 1 所示，在天空可视度良好的情况下，视觉惯性里程计、GNSS载波相位和GNSS伪距均可感知该轨迹。我们可以观察到，GNSS 伪距比另外两种模态的噪声大得多，相对于另外两种模态的

[1]我们的算法是随意的，与扩展卡尔曼滤波器相同。



(c) 我们的轨迹与真实值的对比；(a) 开阔天空下的绝对位姿误差，遮挡 10 颗卫星，平均剩余 4 颗卫星。RTKLIB 在部分历元失效；(b) 轻度遮挡区域的绝对位姿误差，遮挡 5 颗卫星，平均有 5 颗卫星可见。RTKLIB 的轨迹，以及轻度遮挡区域中 ECEF 坐标系下的真实值

图 3：城市峡谷中模拟信号遮挡时与真实值的性能对比。建议放大查看。

定位误差高达 10 米。这表明，将视觉惯性信息与 GNSS伪距融合并不合适，因为视觉惯性里程计无法通过与 GNSS伪距定位对齐，实现相对于 ECEF 坐标系的厘米级配准。GVINS [7]，表 II 中报告的 ECEF配准误差也支持这一结论：即使在开阔天空下的运动场上，配准误差仍达到米级。

	开阔天空，10 颗卫星被遮挡	轻度遮挡，5 颗卫星被遮挡
	RTKLIB：我们的方法	RTKLIB：我们的方法
均方根误差（m）↓	0.68 :**0.38**	0.78 :**0.033**
均值（m）↓	0.55 :**0.34**	0.45 :**0.025**
中位数（m）↓	0.40 :**0.32**	0.26 :**0.019**
标准差（m）↓	0.41 :**0.17**	0.63：**0.023**
最大值 (m) ↓	2.97 :**0.88**	4.90 : **0.43**
失败率 (%) ↑	10.8 :**20.0**	48.6 : **92.9**

表 II：真实值对比实验的统计数据。“10 sats ex”表示手动排除了 10 颗卫星。“5 sates ex”同理。

C. 真值对比实验

为了评估我们的方法在城市场景中遭遇严重天空遮挡时的性能，我们需要创造性地获取真值。将融合低精度伪距 GNSS 测量的 VIO 方法与我们的方案进行比较也不合适，因为这类方法无法融合载波相位 GNSS 测量。因此，我们直接采用载波相位定位结果作为真值。为此，我们采集了两个数据集，一个来自开阔天空环境，另一个来自树木造成轻度遮挡的环境；随后手动排除卫星以模拟天空遮挡，并使用 RTKLIB 处理全星座数据，将结果作为真值。

我们运行所提出的方法和 RTKLIB，在手动移除卫星后对部分卫星测量值进行处理（开阔天空数据集排除了 10 颗卫星，轻度遮挡数据集排除了 5 颗卫星）。被移除的卫星包括观测期间最常被观测到的卫星以及观测频率中度的卫星。图 3 使用 evo 工具包 [27]绘制了随时间变化的绝对位置误差（APE）。均方根误差（RMSE）、平均绝对位姿误差（mean APE）、绝对位姿误差中位数（median APE）、绝对位置误差标准差、最大绝对位姿误差（max APE）以及固定解率（FSR，越高越好）列于表 II 中。对于 RTKLIB 和我们的方法，我们对整数解验证采用相同的标准，即常用的比值检验 [23]，：次优残差与最优残差之比为 3。可以观察到，与 RTKLIB 相比，我们的方法实现了更低的位置误差，同时固定解率（FSR）显著更高；这里的比较基准是真值。



图6：在谷歌地球中显示的图5b轨迹所经城市峡谷。该峡谷的平均建筑物高度为55米，最高建筑物高81米。

D. 真实城市峡谷实验

我们在图4和图5所示的两个具有挑战性的城市峡谷环境“UC1”和“UC2”中开展真实场景实验。 [2] 数据集中的城市峡谷三维街景如图6所示，以说明我们问题的挑战性。正如V-B节所分析的，融合GNSS伪距和视觉惯性里程计（VIO）的工作不适合用于我们的实验对比，因为由于伪距定位的噪声，其向地心地固坐标系配准存在米级偏差。此外，这些方法也没有像我们一样解决载波相位定位问题。

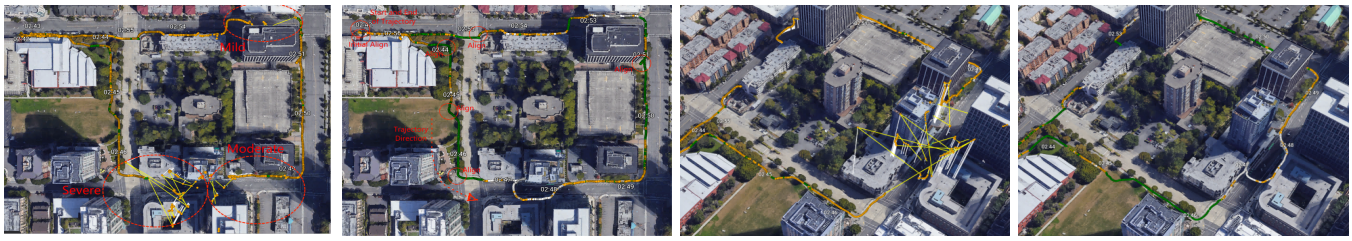
轨迹中，绿色表示固定解，黄色表示浮点解，白色表示在算法1中的验证失败时由对齐后的视觉惯性里程计补出的“填补”位置。图4和图5展示了RTKLIB与我们模型的结果可视化。可以观察到，RTKLIB在水平方向（图4a、5a）和垂直方向（图4c、5c）都受到城市峡谷效应的影响。所有的锯齿状波动

[2]读者还可访问 https://youtu.be/PCDymbSTndQ 观看视频演示。



(a) UC1 中的 RTKLIB (b) 我们的模型在 UC1 中 (c) RTKLIB，UC1（放大视图）(d) 我们的方法，UC1（放大视图）)

图 4：UC1 环境。绿色：固定解。黄色：浮点解。没有有效浮点解时为白色：对齐后的视觉惯性里程计。建议以彩色放大查看。



(a) UC2 中的 RTKLIB (b) 我们的模型在 UC2 中 (c) RTKLIB，UC2（放大视图）(d) 我们的方法，UC2（放大视图）

图 5：UC2 环境。绿色：固定解。黄色：浮点解。没有有效浮点解时，白色表示对齐后的视觉惯性里程计。最佳观看效果请使用彩色显示并放大。

以及RTKLIB的扩展卡尔曼滤波器造成的轨迹大幅跳变，都被我们的模型消除了（图4b、5b）。此外，白色点构成的路段与由绿色和黄色点构成的GNSS定位路段平滑衔接，这从定性上表明我们的融合方法具有连贯性。

为进行定量评估，我们分别从 UC1 和 UC2 轨迹中选取了三个路段，这些路段的城市峡谷效应分别为重度、中度和温和，如图 4a 和图 5a 所示。由于所选路段缺少真实值，我们采用本文方法作为参考，评估 RTKLIB 和视觉惯性里程计相对于本文方法的偏差。[3] 通过相邻路段对齐视觉惯性里程计轨迹。结果表明，尽管视觉惯性里程计的漂移率约为 0。5%，其结果仍比 RTKLIB 的扩展卡尔曼滤波器更接近本文方法的轨迹，这说明我们的系统比原始 RTKLIB 稳健得多。均方根误差、平均绝对位姿误差、绝对位姿误差中位数和最大绝对位姿误差列于以下两张表中。为简便起见，表中将 RTKLIB 记为“RTK”，将视觉惯性里程计记为“VIO”。

UC1 路段	重度 实时动态定位：视觉惯性里程计	中度 实时动态定位：视觉惯性里程计	Mild 实时动态定位：视觉惯性里程计
均方根误差（米）↓	45.39：**1.61**	9.59：**0.99**	6.58：**1.11**
均值（米）↓	32.08：**1.29**	4.79：**0.69**	4.42：**1.02**
中位数（m）↓	17.00：**0.96**	1.32：**0.31**	1.14：**0.94**
最大值（米）↓	82.06：**4.15**	51.16：**3.30**	14.94：**1.85**

[3]RTKLIB 通常被用作 SLAM 研究中的真实值。就此而言，我们改进了许多研究所使用的真实值。

UC2 分段	重度 实时动态定位：视觉惯性里程计	中度 实时动态定位：视觉惯性里程计	Mild 实时动态定位：视觉惯性里程计
均方根误差（米）↓	44.29：**2.33**	23.21：**1.55**	9.51：**0.79**
均值 (m) ↓	37.51：**1.88**	14.69：**1.11**	9.38：**0.63**
中位数 (m) ↓	38.14：**1.02**	5.57：**0.87**	8.59：**0.59**
最大值 (m) ↓	112.44：**4.74**	59.27：**4.10**	12.40：**1.91**

最后，我们计算了 RTKLIB 的扩展卡尔曼滤波器和我们的方法的固定解率，结果如下表所示。我们采用与第 V-C 节相同的整数解验证比值检验标准。结果表明，我们的方法显著提高了固定解率（越高越好），这意味着我们的方法在城市峡谷中具有更强的鲁棒性。

环境	UC1 实时动态定位：我们的方法	UC2 实时动态定位：我们的方法
固定解率 (%) ↑	39.4：**50.8**	8.1：**36.5**

VI. 结论

本文针对城市峡谷中全球导航卫星系统载波相位定位的鲁棒性问题，采用松耦合扩展卡尔曼滤波器融合全球导航卫星系统载波相位与视觉惯性里程计。我们提出一种基于优化的方法，用于标定惯性测量单元与全球导航卫星系统天线之间的外参，从而将视觉惯性里程计准确对齐至地心地固坐标系。我们使用视觉惯性里程计作为扩展卡尔曼滤波器的预测步骤，并将其协方差估计纳入扩展卡尔曼滤波器的协方差传播中。我们通过仿真实验以及在复杂城市峡谷中的真实场景实验，对所提方法的性能进行了定量验证。在复杂城市峡谷中，该方法的鲁棒性、定位精度和固定解率均大幅优于最先进的GNSS载波相位定位方法 RTKLIB [10]，弥补了实现普适全球定位方面的不足。

参考文献

[1] A. I. Mourikis 和 S. I. Roumeliotis，“用于视觉辅助惯性导航的多状态约束卡尔曼滤波器”，载于《2007年IEEE机器人与自动化国际会议（ICRA）论文集》。IEEE，2007年，第3565–3572页。[2] T. Qin、P. Li 和 S. Shen，“VINS-Mono：一种鲁棒且通用的单目视觉惯性状态估计器”，《IEEE机器人学汇刊》，第34卷，第4期，第1004–1020页，2018年。[3] C. Campos、R. Elvira、J. J. G. Rodr´ıguez、J. M. Montiel 和 J. D. Tard´os，“ORB-SLAM3：用于视觉、视觉惯性和多地图SLAM的精确开源库”，《IEEE机器人学汇刊》，第37卷，第6期，第1874–1890页，2021年。[4] G. Cioffi 和 D. Scaramuzza，“基于优化的视觉惯性里程计中全局位置测量值的紧耦合融合”，载于《2020年IEEE/RSJ智能机器人与系统国际会议（IROS）论文集》。IEEE，2020年，第5089–5095页。[5] W. Lee、K. Eckenhoff、P. Geneva 和 G. Huang，“间歇性GPS辅助视觉惯性里程计：在线初始化与标定”，载于《2020年IEEE机器人与自动化国际会议（ICRA）论文集》。IEEE，2020年，第5724–5731页。[6] T. Qin、S. Cao、J. Pan 和 S. Shen，“一种用于多传感器全局位姿估计的通用优化框架”，arXiv预印本 arXiv:1901.03642，2019年。[7] S. Cao、X. Lu 和 S. Shen，“GVINS：用于平滑且一致状态估计的紧耦合GNSS-视觉-惯性融合”，《IEEE机器人学汇刊》，2022年。[8] S. Boche、X. Zuo、S. Schaefer 和 S. Leutenegger，“采用紧耦合抗丢失GPS融合的视觉惯性SLAM”，arXiv预印本 arXiv:2208.00709，2022年。[9] “Project Aria”，https://about.facebook.com/realitylabs/projectaria/。[10] T. Takasu 和 A. Yasuda，“采用开源程序包RTKLIB开发低成本RTK-GPS接收机”，载于《GPS/GNSS国际研讨会》，第1卷，2009年。[11] T. Li 、 H. Zhang 、 Z. Gao 、 X. Niu 和 N. El-Sheimy，“ 融合单目相机、MEMS-IMU与单频多GNSS实时动态定位，以实现在GNSS受限环境中的精密导航”，《Remote Sensing》，第11卷，第6期，第610页，2019年。[12] J. Liu、W. Gao 和 Z. Hu，“与原始GNSS测量值紧耦合的基于优化的视觉惯性SLAM”，载于《2021年IEEE机器人与自动化国际会议（ICRA）论文集》。IEEE，2021年，第11 612–11 618页。[13]J. E. Yoder、P. A. Iannucci、L. Narula 和 T. E. Humphreys，“用于微型飞行器位姿估计的多天线视觉与惯性辅助CDGNSS”，载于《无线电导航实验室会议论文集》，2020年。[14] R. Mur-Artal 和 J. D. Tard´os，“ORB-SLAM2：用于单目、立体和RGB-D相机的开源SLAM系统”，《IEEE机器人学汇刊》，第33卷，第5期，第1255–1262页，2017年。[15] Y. Yu、W. Gao、C. Liu、S. Shen 和 M. Liu，“普适环境中的GPS辅助全向视觉惯性状态估计器”，载于《2019年IEEE/RSJ智能机器人与系统国际会议（IROS）论文集》。IEEE，2019年，第7750–7755页。[16] R. Mascaro、L. Teixeira、T. Hinzmann、R. Siegwart 和 M. Chli，“GOMSF：用于鲁棒无人机位姿估计的图优化多传感器融合”，载于《2018年IEEE机器人与自动化国际会议（ICRA）论文集》。IEEE，2018年，第1421–1428页。[17] D. P. Shepard 和 T. E. Humphreys，“通过融合视觉SLAM、载波相位GPS和惯性测量值，实现高精度全球参考位置与姿态”，载于《2014年IEEE/ION定位、位置与导航研讨会（PLANS 2014）》。IEEE，2014年，第1309–1328页。[18] T. Takasu 和 A. Yasuda，“通过MEMS-IMU/GPS集成检测并修复周跳，适用于移动环境RTK-GPS”，载于《国际导航学会卫星部第21届国际技术会议（ION GNSS）论文集》，2008年，第64–71页。[19] W. Wen、G. Zhang 和 L.-T. Hsu，“利用三维激光雷达和建筑物高度校正GNSS非视距”，载于《国际导航学会卫星部第31届国际技术会议（ION GNSS+）论文集》，2018年，第3156–3168页。[20] P. Misra 和 P. Enge，《全球定位系统：信号、测量与性能》，Ganga-Jamuna Press，纽约，2006年。[21] P. Teunissen，“最小二乘模糊度去相关调整：一种快速GPS整数模糊度估计方法”，《Journal of Geodesy》70，第65–82页，1995年。[22] X.-W. Chang、X. Yang 和 T. Zhou，《Mlambda：一种用于整数最小二乘估计的改进 lambda 方法》，《*Journal of Geodesy*》，第 79 卷，第 9 期，第 552–565 页，2005 年。[23] S. Verhagen 和 P. J. Teunissen，《面向未来 GNSS 模糊度解算的比值检验》，《*GPS*解决方案》，第 17 卷，548，2013 4，535 no. pp. – 。[24] M. Dahleh，《系统辨识讲义，MIT 6.345》，2005 年。[25] H. Hotelling，《将一组统计变量分解为主成分》，《教育心理学杂志》，第 24 卷，第 6 期，第 417 页，1933 年。[26] J. Sola、J. Deray 和 D. Atchuthan，《机器人学状态估计的微型李群理论》，*arXiv* 预印本 *arXiv:1812.01537*，2018年。[27] 《用于评估里程计和 SLAM 的 Python 软件包》，https:// github.com/MichaelGrupp/evo。