

Leader-Follower cooperative localization based on VIO / UWB loose coupling for AGV group

Haifeng Zhang^{1,2}, Zhitian Li¹, Shuaikang Zheng^{1,2}, Yunfei Liu^{1,2}, Pengcheng Zheng^{1,2} and Xudong Zou^{1*,2,3}

*Correspondence: zouxd@aircas.ac.cn

¹The State Key Laboratory of Transducer Technology, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, 100010, CHINA.

²School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing, 100010, CHINA.

³QiLu Research Institute, Aerospace Information Research Institute Chinese Academy of Sciences, Jinan, 250000, CHINA.

Commission IV, WG IV/5

KEY WORDS: Cooperative localization, Leader-Follower, VIO, UWB, Graph optimization.

ABSTRACT:

In this paper, we propose a *Leader-Follower* cooperative localization system for AGV group, which equipped with camera/IMU/UWB anchors as the *Leader* and IMU/UWB tags as the *Follower*. Firstly, we introduce the plane constraint with SE(2) into VIO to suppress drift in 2D motion; Secondly, the *Follower* receives the *Leader's* VIO pose and calculates the relative pose through the multiple groups of UWB ranging. The algorithm of *Leader* and *Follower* is decoupled, which means *Leader* can choose different self-localization technology to adapt to different scenario, such as LiDAR-Inertial Odometry(LIO), GNSS, etc. In the *Follower*, a tightly coupled graph optimization based algorithm to fuse UWB ranging and IMU is proposed to estimate its relative pose to the *Leader*. Lastly, extensive field experiment shows the system can realize omni-directional cooperative positioning in indoor and outdoor environments and the accuracy of collaborative positioning reaches centimeter level.

1. INTRODUCTION AND RELATED WORK

In the active field of robot localization, multi-robot cooperative localization(CL) in GNSS denied environments has always been an attractive research. Its purpose is to express the position and orientation of the robots in a common frame. The motivation of this research comes from swarm detection(Michael, 2013), collaborative mapping(Nguyen, 2017), formation control(Turpin, 2012) and so on. (Saedi, 2016)gives a comprehensive review of multi-robot SLAM technologies and points out one most key issue: relative pose estimation. There are some sensors can get relative measurements, such as: cameras and LiDARs for common features, UWB for relative distances, etc.

According to the type of sensor, AGVs cooperative localization can be divided into feature-based(Danping ZOU, 2019) and range-based(Buehrer, 2018). (Dube, 2017)presented a cooperative localization framework of multiple AGV in buildings based on 3D LiDAR, which relies on inter-robot place recognition to estimate the relative pose between AGVs. Vision for Robotics Lab, ETH Zurich(Schmuck, 2021) proposes a centralized localization system, in which each node runs VIO and uploads the feature descriptor of the image to the central processor, and rely on the closed-loop detection algorithm to estimate the relative state between nodes. The impressive absolute trajectory error(ATE) smaller than 5cm is reported in this paper, which shows the strong advantage of VIO algorithm in cooperative localization. However, both schemes have some disadvantages. First, cameras or LiDAR are limited by the field of view(FOV) and cannot achieve omni-directional relative localization. Secondly, LiDAR and camera' data occupy large network bandwidth and more computing power. (Xu, H, 2020) proposed a decentralized collaborative localization method. Each node calculates the relative pose through target detection and tracking. Although the limitation of FOV is overcome by using fisheye camera, there is still a problem of robustness in target tracking.

Ranging based method has attracted many researchers. (Jung, 2021) proposes a distributed cooperative location algorithm

based on filtering, through the sharing of approximate covariance between AGVs. The ranging information comes from the fixed and calibrated UWB anchors, and the decimeter-level positioning accuracy is reported in this paper. However, fixed UWB anchors are sometimes difficult to deploy in outdoor environments. There are some co-location work that integrates VIO+UWB and achieves high accuracy(Queralt, 2021). A novel idea is to install UWB on different AGVs to form a mobile anchor system and get ranging information between agents. All AGVs need to be equipped with cameras, IMU and UWB. It is necessary to consider a simpler configuration, which is very important when the agent computing resources are limited.

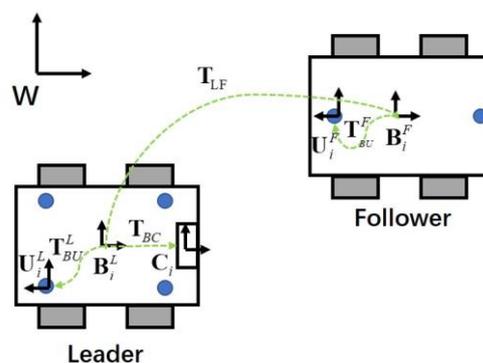


Figure 1. The frame definition of the *Leader* and *Follower*: the world frame is W ; the AGV base frame (IMU frame) is B ; the camera frame is C ; the UWB frame is U ; the transformation of *Leader* and *Follower* is T_{LF} .

Another very important topic is the state estimation algorithm, which can be divided into filtering-based(Jazwinski, 1970) and graph optimization based(Qin,T, 2019). Filtering-based method is recursive and sensitive to sensor time consistency, due to its first-order Markov process assumption. Graph optimization is a

batch optimization algorithm that usually maintains a sliding window(Demmel, 2021). This algorithm is widely used for robot state estimation and achieves excellent performance. In sliding window, all sensor data and the state are coupled in the form of cost function and solved by nonlinear optimization method, such as, GN or LM algorithm, so it's less strict about temporal consistency. Although, its cost of calculation is generally larger than the filtering method. There is common data delay in the communication network for multi-AGV system, so graph optimization algorithm is more suitable for multi-robot problem.

In this paper, we propose a cooperative localization system for AGV group, which equipped with camera/IMU/UWB anchors as the *Leader* and IMU/ UWB tags as the *Follower*. Different from the centralized system, we solve the pose locally on each agent's processor. Visual-Inertial odometry(Qin, T, 2018) (VIO) is an excellent state estimation technology for single robot, which is mainly used for self-localization of *Leader*. However, in 2D motion scene, IMU usually produces large cumulative error in Z axis due to insufficient excitation. We add the plane constraint(Zheng, F, 2019) to the original VIO pose to suppress the cumulative error. This is a soft constraint that fully takes into account the fluctuations of the ground. Unlike their work, we do not parameterize the pose on SE(2), but directly on SE(3), which allows us to directly use the results of VIO. In order to ensure the robustness of the system and successful initialization, four UWB anchors are equipped on the *Leader*. Assuming that the UWB anchors are rigidly connected to AGV body, which eliminates the restriction of fixed anchors(Li, J, 2018).

Followers calculate the relative pose to *Leader* through UWB ranging(Nguyen, 2018), equip at least 2 UWB tags to ensure the observability of the orientation and use IMU for state propagation. Then, we propose a tightly coupled graph optimization algorithm to realize the relative localization between *Leader* and *Followers*. UWB and IMU provide ranging factor and preintegration(Forster, 2015) factor respectively. It should be emphasized that, unlike VIO+UWB methods above, there is no camera on *Followers*, which can greatly reduce the amount of computation and hardware resources, and simpler configuration brings greater challenges to our systems and algorithms.

As mentioned above, the existing AGV cooperative localization methods have some shortcomings: UWB-based methods generally rely on fixed anchors and have low accuracy; VIO+UWB based anchor-free methods have higher accuracy, but no consider simpler configuration. The method proposed in this paper adopts a simpler configuration and achieves centimeter-level positioning accuracy. The graph optimization algorithm of tightly coupled UWB+IMU is the main highlight of this paper. We summarize the contributions of this paper as follows:

- (1) A *Leader-Follower* cooperative localization system based on VIO+UWB is proposed for AGV group. We adopt a simpler configuration and suppress the drift of VIO through plane constraints.
- (2) On the *Follower*, a tightly coupled localization algorithm based on graph optimization is proposed, which fuses the UWB ranging factor and the IMU pre-integration factor.
- (3) Extensive real-world experiments show the advanced nature of our system, reaching centimeter-level accuracy in line-of-sight environments.

The remainder of this paper is organized as follows. 2 describes the meaning of the symbols used in this paper, and 3 describes proposed system and algorithm, including sensor time synchronization, *Leader* and *Followers* localization algorithms. Experiments are described in detail in Section 4, including *Leader* and AGV group localization experiments. and analysis

of positioning errors. Finally, we analyze the experimental data and make a conclusion.

2. PRELIMINARIES

Throughout this manuscript, \mathbf{T}_{AB} denotes the transformation that maps the coordinates in any frame $\{A\}$ to $\{B\}$. \mathbf{R}_{AB} and \mathbf{p}_{AB} denote the rotation and translation respectively. $\mathbf{R} \in \text{SO}(3)$ and $\mathbf{p} \in \mathbf{R}^3$. The expected state estimation based on current measurement is χ^* . Considering time domain, $\{A_t^x\}$ denotes frame $\{A\}$ of AGV \mathbf{X} at instant t . Lie algebras follow the same naming pattern. The frame used in this paper is shown in Figure 1.

we define the state of each agent as χ . It is equivalent to $[\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n]$ where, n is the size of the window, and $\mathbf{x}_k = [\mathbf{p}_t, \mathbf{v}_t, \mathbf{q}_t, \mathbf{b}_a, \mathbf{b}_w]$, $k \in [0, n]$, respectively denote position, speed, orientation in the world frame and IMU bias.

3. LEADER-FOLLOWER COOPERATIVE LOCALIZATION SYSTEM AND ALGORITHM

We divide the AGV group into a *Leader* and several *Followers*. The *Leader* installed four UWB anchors and a stereo VI sensor. The coordinates of UWB relative to the AGV body are calibrated in advance. *Leader* establishes a common reference frame(world frame) for all agents. The *Follower* installed UWB tags on the front and rear of AGV, and installed IMU in the middle of the AGV body, which is used for pre-integration between two UWB ranging. Figure 1 shows the installation of the sensors on AGV.

Communication between agents relies on WiFi. The *Leader* broadcasts the improved VIO pose. On the *Followers*, UWB ranging and IMU input to relative state estimator, and converts to the *Leader's* frame (world frame).The detailed system framework is shown in Figure 2.

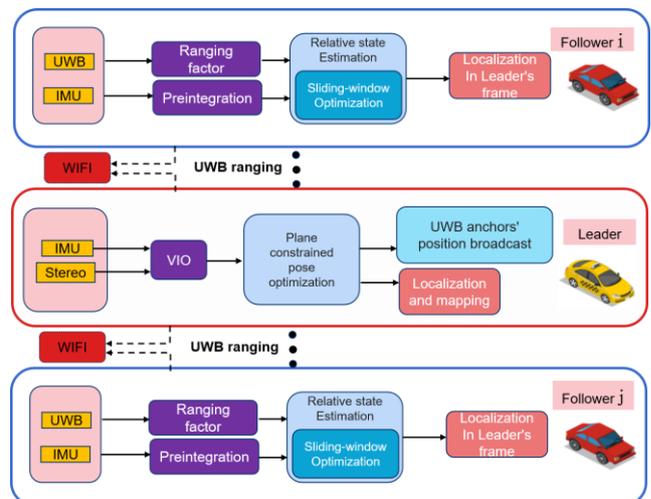


Figure 2. Framework of *Leader-Follower* cooperative localization system. The *Leader* runs VIO for self-localization, assuming that there is a rigid connection between UWB anchors and AGV body, and broadcasts the pose through WiFi. The *Follower* run the tightly coupled graph optimization algorithm for relative state estimation.

On each *Follower*, we define the maximum posterior estimation (MPE) problem of the system, and the estimated state is and *Follower's* state and the pose of *Leader*. That is, the final objective function to be solved is:

$$\begin{aligned} \chi^* &= \arg \max_{\chi, \mathbf{T}_{WB}^L} p(\chi) \prod_{i=1}^n p(\mathbf{z}_i | \chi) \\ &= \arg \min_{\chi, \mathbf{T}_{WB}^L} \sum_i \left\{ \|\mathbf{r}(\mathbf{z}_{imu}, \chi)\|^2 + \|\mathbf{r}(\mathbf{z}_{uwb}, \chi)\|^2 \right\} \end{aligned} \quad (1)$$

where \mathbf{z} represents the sensor measurement and $\mathbf{r}(\cdot)$ denotes the residual function. As shown in (1), the *Leader's* pose has been optimized on all agents, and we take the average of all results as the *Leader's* final pose.

$$\mathbf{T}_{WB}^L = \frac{\sum_{i=1}^m \mathbf{T}_{WB}^{Li}}{m} \quad (2)$$

m is the number of agents.

3.1 Data synchronization

The system needs to synchronize data from three different types: VIO, UWB, and IMU. They have different frequencies: VIO has the lowest frequency; IMU has the highest frequency. Time synchronization of these data is a challenging problem. To maximize the utilization of all sensor data, we propose a UWB-centric approach to data synchronization. See **Figure 3**. We adopt first-order linear interpolation between VIO and UWB. It is assumed that the *Leader* is moving at uniform speed during Δt , usually it is less than 0.05s. We preintegrate the IMU data between two UWB ranging.

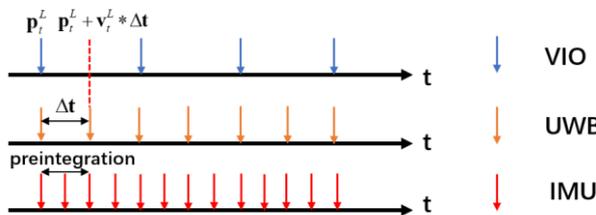


Figure 3. We use first-order linear interpolation between VIO and UWB and preintegrate the IMU between UWB ranging.

3.2 Leader: self_localization based on VIO with Planar Constraints

AGVs move on a two-dimensional plane, which means that there is no excitation source in the Z-axis direction, and it results in drift in the Z-axis, which can be coupled to the position of the UWB anchors and affect the localization performance of the system. Therefore, we propose a SE(2)-constraint to suppress the drift in Z-axis.

It should be noted that we did not add the deterministic SE(2)-constraint, because considering the undulation of the outdoor ground. We added this constraint to VIO's pose in SE(3), instead of parameterizing it with SE(2). This allows us to directly utilize the output of the VIO.

we assume that AGV moves in a 2D plane, so its pose can be expressed as $se(3)$ in Lie algebra:

$$\varphi = [0, 0, w_3, v_1, v_2, 0]^T, \varphi \in se(3) \quad (3)$$

Where w_3 is related to the yaw angle of AGV and v_1, v_2 represents the position on the plane. Then we construct a residual and defined as a nonlinear least squares problem:

$$\mathbf{r}_s = \exp(\varphi^\wedge)^{-1} \mathbf{T}_{WB}^L \quad (4.1)$$

$$\mathbf{T}_{WB}^{L*} = \underset{\mathbf{T}_{WB}^L}{\operatorname{argmin}} \|\mathbf{r}_s\|^2 \quad (4.2)$$

Where $\exp(\bullet)$ represents the exponential mapping from $se(3)$ to SE(3). The covariance matrix of this observation can be expressed as:

$$\Omega = \operatorname{diag} \left(\delta, \delta, \frac{1}{\sigma_w^2}, \frac{1}{\sigma_v^2}, \frac{1}{\sigma_v^2}, \delta \right) \quad (5)$$

δ is close to 0, and σ is related to the pose covariance of the VIO. Assuming that the UWB anchors on the *Leader* is rigidly connected to the body, the UWB anchors' position in the world frame can be expressed as:

$$\mathbf{p}_{WU^L} = \mathbf{p}_{WB^L} + \mathbf{R}_{WB^L} \mathbf{p}_{B^L U} + \eta_{vio} \quad (6)$$

η_{vio} is white Gaussian noise.

3.3 Follower: graph optimization based UWB + IMU tight coupling localization

For each *Follower*, we solve the multi-observation-based maximum likelihood estimation(MLE) problem (1) in a sliding window, which can be modeled as a factor graph; UWB provides ranging factors and IMU provides preintegration factors. The detailed factor graph model is shown in **Figure 4**.

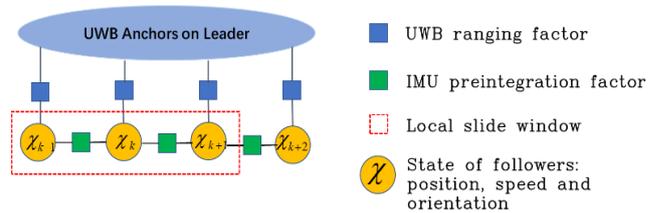


Figure 4. Factor graph model of *Follower*. UWB provides ranging factors and IMU provides preintegration factors.

3.3.1 UWB factor

UWB is a potential ranging technology, which is very suitable for robot localization. The ranging error in the line-of-sight environment is within 10cm. Besides, each UWB device can be assigned a unique ID, which makes it easy to distinguish different robots. We used the TW-TOF protocol, for it does not require strict time synchronization between UWB devices.

Considering the distance between *Leader* and *Followers*, the UWB measurement can be modeled as:

$$\begin{aligned} d &= \|\mathbf{p}_{wu^L} - \mathbf{p}_{wu^F}\| + \eta_{uwb} \\ &= \left\| (\mathbf{p}_{WB^L} + \mathbf{R}_{WB^L} \mathbf{p}_{B^L U}) - (\mathbf{p}_{WB^F} + \mathbf{R}_{WB^F} \mathbf{p}_{B^F U}) \right\| + \eta_{uwb} \end{aligned} \quad (7)$$

Further UWB residual can be expressed as:

$$\mathbf{r}(\mathbf{z}_{uwb}, \boldsymbol{\chi}) = d - \left\| (\mathbf{p}_{WB^L} + \mathbf{R}_{WB^L} \mathbf{p}_{B^L U}) - (\mathbf{p}_{WB^F} + \mathbf{R}_{WB^F} \mathbf{p}_{B^F U}) \right\| \quad (8)$$

We have obtained \mathbf{p}_{WU^L} from (6), and η_{uwb} is Gaussian noise with zero mean. The initial value of $\mathbf{p}_{B^F U}$ is get by trilateration. Although distance-based localization can obtain closed-form solutions through trilateration algorithms, nonlinear optimization methods can achieve higher accuracy through multiple iterative calculations. In addition, the optimized method can also estimate the orientation of the robot, which cannot be calculated by trilateration.

3.3.2 IMU factor

The IMU raw data is angular velocity and linear acceleration with noise, which includes the Earth's gravity component. We ignore the rotation of the earth. IMU measurements can be modeled as:

$$\begin{aligned} \tilde{\mathbf{a}}_t &= \mathbf{a}_t + \mathbf{b}_{a_t} + \mathbf{n}_a \\ \tilde{\boldsymbol{\omega}}_t &= \boldsymbol{\omega}_t + \mathbf{b}_{\omega_t} + \mathbf{n}_\omega \end{aligned} \quad (9)$$

Where $\tilde{\mathbf{a}}_t$ and $\tilde{\boldsymbol{\omega}}_t$ denote the angular velocity and linear acceleration of the output of IMU, \mathbf{a}_t and $\boldsymbol{\omega}_t$ denote true value, and \mathbf{b}_a , \mathbf{b}_ω are the bias of IMU. The noise \mathbf{n}_a and \mathbf{n}_ω are assumed to be zero-mean Gaussian distributed.

In engineering, the frequency of the IMU is usually higher than that of the UWB, and the IMU preintegration[19] method is used to aggregate multiple measurements into one. For inertial measurements in the time interval $[t_k, t_{k+1}]$, the derived measurements are calculated as:

$$\begin{aligned} \boldsymbol{\alpha}_{b_{k+1}}^{b_k} &= \iint_{t \in [t_k, t_{k+1}]} \mathbf{R}_{b_t}^{b_k} (\tilde{\mathbf{a}}_t - \mathbf{b}_{a_t}) dt^2 \\ \boldsymbol{\beta}_{b_{k+1}}^{b_k} &= \int_{t \in [t_k, t_{k+1}]} \mathbf{R}_{b_t}^{b_k} (\tilde{\mathbf{a}}_t - \mathbf{b}_{a_t}) dt \\ \boldsymbol{\gamma}_{b_{k+1}}^{b_k} &= \int_{t \in [t_k, t_{k+1}]} \frac{1}{2} \boldsymbol{\Omega}(\boldsymbol{\omega}_t - \mathbf{b}_{\omega_t}) \boldsymbol{\gamma}_{b_t}^{b_k} dt \end{aligned} \quad , \quad (10)$$

with

$$\boldsymbol{\Omega}(\boldsymbol{\omega}) = \begin{bmatrix} -[\boldsymbol{\omega}]_x & \boldsymbol{\omega} \\ -\boldsymbol{\omega}^T & 0 \end{bmatrix}, [\boldsymbol{\omega}]_x = \begin{bmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{bmatrix} \quad (11)$$

$\{\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}\}$ encapsulates the relative position, velocity and rotation information between two UWB ranging. Finally, the residual and system state can be expressed as:

$$\mathbf{r}_B(\mathbf{z}_{imu}, \boldsymbol{\chi}) = \begin{bmatrix} \mathbf{R}_w^{b_k} \left(\mathbf{p}_{b_{k+1}}^w - \mathbf{p}_{b_k}^w + \frac{1}{2} \mathbf{g}^w \Delta t_k^2 - \mathbf{v}_{b_k}^w \Delta t_k \right) - \boldsymbol{\alpha}_{b_{k+1}}^{b_k} \\ \mathbf{R}_w^{b_k} \left(\mathbf{v}_{b_{k+1}}^w + \mathbf{g}^w \Delta t_k - \mathbf{v}_{b_k}^w \right) - \boldsymbol{\beta}_{b_{k+1}}^{b_k} \\ 2 \left[\mathbf{q}_{b_k}^{w-1} \otimes \mathbf{q}_{b_{k+1}}^w \otimes \left(\boldsymbol{\gamma}_{b_{k+1}}^{b_k} \right)^{-1} \right]_{xyz} \\ \mathbf{b}_{a_{b_{k+1}}} - \mathbf{b}_{a_{b_k}} \\ \mathbf{b}_{\omega_{b_{k+1}}} - \mathbf{b}_{\omega_{b_k}} \end{bmatrix} \quad (12)$$

4. EXPERIMENT

In this section, we did two experiments. The first experimental scenario is the *Leader* localization experiment to verify the suppression of VIO drift by the plane-constraint. We adopt the state-of-the-art stereo inertial odometry(Qin, 2019). In the second experiment, we show the co-location of three AGVs to verify the effectiveness of the proposed system and algorithm. RTK is used as the ground truth in the experiment, because the positioning error of RTK outdoors is less than 3cm.

In the experiments, we adopt the configuration shown in **Figure 5**. On the *Leader*, realsenseD435i VI kit provide stereo + IMU; UWB model using DW1000; RTK using G200; processor is NVIDIA joston AGX. On *Followers*, using the same UWB and IMU is using CH110.

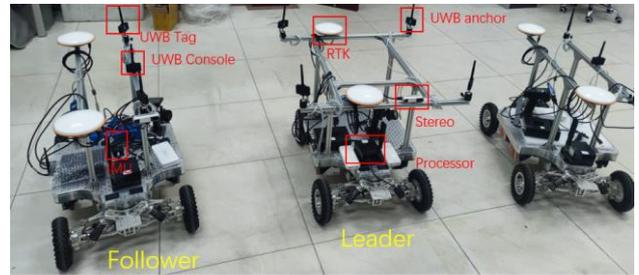


Figure 5. Sensor configuration of AGV group

4.1 Plane Constraint Effectiveness Experiment

In order to suppress the drift (mainly in the Z axis) of the *Leader's* motion in the 2D plane, we add the SE(2)-plane constraint to its VIO trajectory (4.1). We did extensive outdoor experiments for *Leader*. We use RTK to provide drift-free reference values. As shown in **Figure 6**.

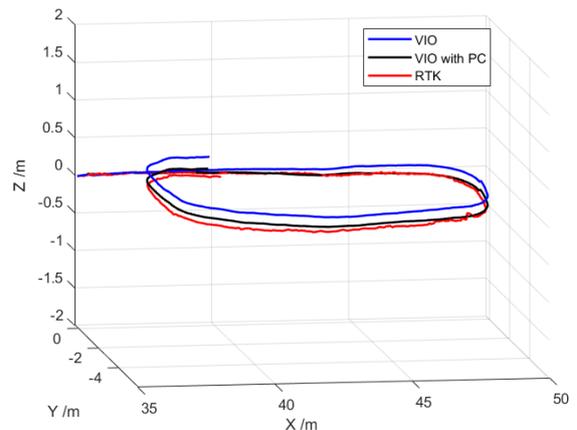


Figure 6. Comparison with and without planar constraints. RTK provided drift-free reference values.

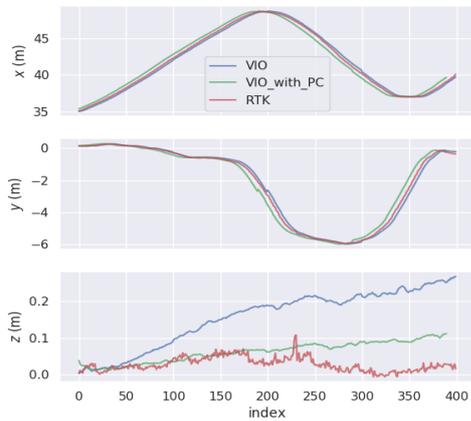


Figure 7. Components of three trajectories in coordinate axes.

Figure 7 shows that the SE(2)-constraint effectively suppresses the 50% drift in the Z axis. It is a soft constraint, so it retains the oscillation on the Z axis, which is mainly caused by ground fluctuations or vibration of the AGV body.

4.2 AGV group Real-world Experiment

In the outdoor scene, we did three AGVs co-location experiments. They consist of a *Leader* and two *Followers*. To ensure time synchronization between AGVs, we use Network Time Protocol (NTP), and each *Follower* is referenced to the *Leader's* clock. The *Leader* kept low speed to move the UWB anchors slowly. *Followers* are distributed on both sides of the *Leader* and maintain a good line-of-sight environment. We install RTK on each AGV to record the real trajectory as ground truth. **Figure 8** shows the trajectories of all AGVs. (Only focus on 2D trajectories).

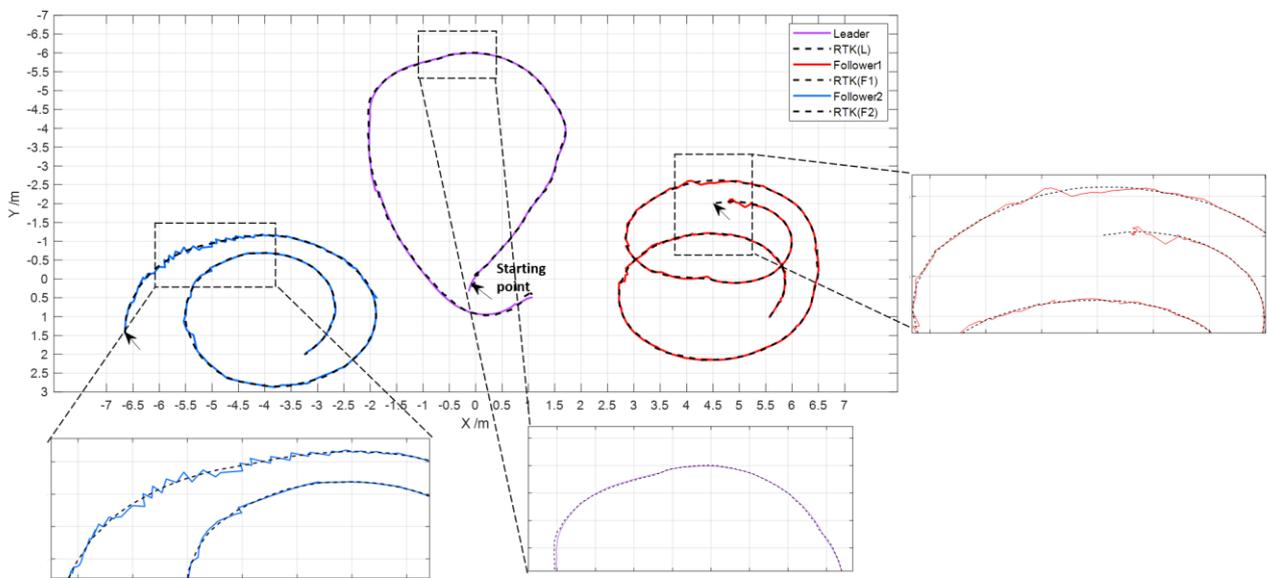


Figure 8. Co-localization experiment of three AGVs, including a *Leader* and two *Followers*. The origin of world frame is the starting point of *Leader*. We zoomed in on the details to better show the experimental results.

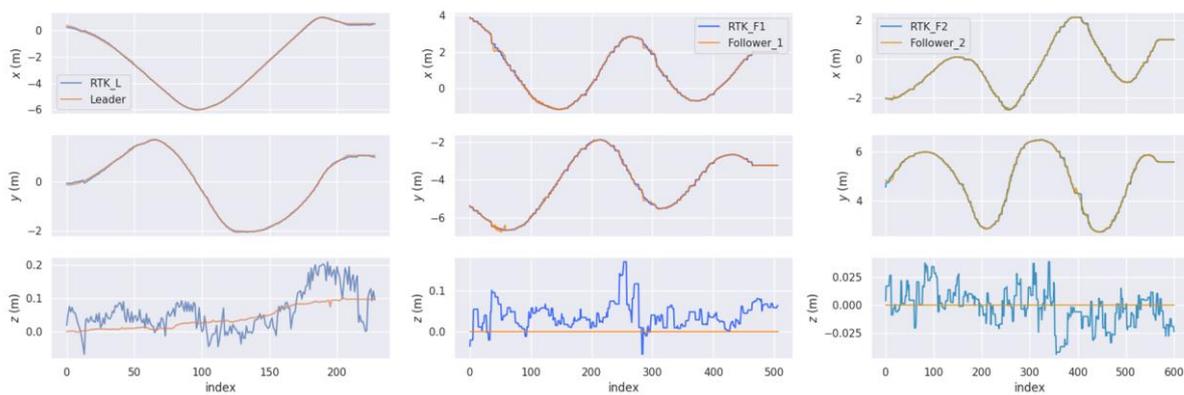


Figure 9. The trajectory of the AGVs group on the axes, in which the *Followers' Z*-axis data is normalized to 0.

The co-location experiment was carried out in a rectangular area with area of 140 m². *Followers* are distributed on both sides of the *Leader* and perform a spiral circular motion, which includes

translation and rotation. Meanwhile, a point to note in experiments is that, *Leader* needs to move slowly because the position of UWB anchors is critical to the localization of AGVs.

The trajectory length of the AGVs exceeded 210 m. The *Followers'* position occasionally fluctuates for two reasons. One is that the UWB tag and AGV body are not rigidly connected when rotating at high angular velocity. The second is the inevitable non-line-of-sight(NLOS) measurement error. For a more detailed study, we analyzed the data of AGVs on the X and Y axis. Because all UWB and camera are at the same height in the experiment, we don't pay attention to the *Followers'* Z-axis data and normalize it to 0. (See **Figure 9**.)

We used the evo tool(MichaelGrupp, 2017) to align the AGVs' trajectory with the ground truth provided by RTK, and calculated **RMSE** and **Maximum error** of position for all AGVs. See **Table 1**.

AGV	RMSE /m	Max error /m
<i>Leader</i>	0.069	0.17
<i>Follower1</i>	0.078	0.31
<i>Follower2</i>	0.095	0.39
Average	0.082	

Table 1. The **RMSE** and **Maximum error** of position for AGVs.

From **Table 1**, it can be inferred that the proposed co-location system and algorithm have good performance. All AGVs' positioning error have reached the centimeter-level. Compared with VIO+UWB-based methods, we achieve the same state-of-the-art results under simpler configuration conditions. The NLOS error of UWB ranging in the experiment results in the *Follower's* **Max error** being greater than that of the *Leader*.

5. CONCLUSION

In this paper, we propose a *Leader-Follower* cooperative localization system and algorithm. This approach does not require a centralized processor. Each agent solves the global pose locally. We suppress the drift of VIO in the Z axis by plane-constraint and propose a co-localization algorithm based on graph optimization. Experiments show that the algorithm achieves centimeter-level positioning accuracy on the AGVs group. It shows competitive performance in positioning accuracy. This system is expected to be applied to multi-robot tasks such as group collaborative detection, collaborative mapping, and rescue. In the future, the non-line-of-sight error of UWB will be fully considered and the robustness of the system will be improved.

REFERENCES

Buehrer, R. M., Wymeersch, H., & Vaghefi, R. M. (2018). Collaborative Sensor Network Localization: Algorithms and Practical Issues. *Proceedings of the IEEE*, 1089-1114.

Danping ZOU, Ping TAN, Wenxian YU. Collaborative visual SLAM for multiple agents: A brief survey. *Virtual Reality & Intelligent Hardware*, 2019, 1(5): 461—482 DOI: 10.1016/j.vrih.2019.09.002.

Dube, R., Gawel, A., Sommer, H., Nieto, J., & Cadena, C. (2017). An online multi-robot SLAM system for 3D LiDARs. 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS).

Demmel, N., Schubert, D., Sommer, C., Cremers, D., & Usenko, V. (2021). Square Root Marginalization for Sliding-Window Bundle Adjustment. In *Proceedings of the IEEE/CVF*

International Conference on Computer Vision (pp. 13260-13268).

Forster, C., Carlone, L., Dellaert, F., & Scaramuzza, D. (2015). IMU Preintegration on Manifold for Efficient Visual-Inertial Maximum-a-Posteriori Estimation. *Georgia Institute of Technology*.

Grupp, M. "evo: Python package for the evaluation of odometry and SLAM," <https://github.com/MichaelGrupp/evo>, 2017.

Jung, R., & Weiss, S. . (2021). Modular multi-sensor fusion: a collaborative state estimation perspective. *IEEE Robotics and Automation Letters*, PP(99), 1-1

Jazwinski, A. H., *Stochastic Processes and Filtering Theory*, New York, Academic, 1970.

Li, J., Bi, Y., Li, K., Wang, K., Lin, F., & Chen, B. M. (2018). Accurate 3D Localization for MAV Swarms by UWB and IMU Fusion. 2018 IEEE 14th International Conference on Control and Automation (ICCA).

Nguyen, T. M., Zaini, A. H., Chen, W., Guo, K., & Xie, L. (2018). Robust Target-Relative Localization with Ultra-Wideband Ranging and Communication. 2018 International Conference on Robotics and Automation (ICRA 2018),

Queralta, J.P., Qingqing, L., Schiano, F., & Westerlund, T. (2020). VIO-UWB-Based Collaborative Localization and Dense Scene Reconstruction within Heterogeneous Multi-Robot Systems. *ArXiv*, abs/2011.00830.

Qin, T., Cao, S., Pan, J., & Shen, S. (2019). A general optimization-based framework for global pose estimation with multiple sensors. *arXiv preprint arXiv:1901.03642*.

Qin, T., Peiliang, Li, Shaojie, & Shen. (2018). VINS-Mono: A Robust and Versatile Monocular Visual-Inertial State Estimator. *IEEE TRANSACTIONS ON ROBOTICS*.

Qin, T., Pan, J., Cao, S., & Shen, S. (2019). A general optimization-based framework for local odometry estimation with multiple sensors. *arXiv preprint arXiv:1901.03638*.

Saeedi, Sajad, Trentini, Michael, Seto, & Mae, et al. (2016). Multiple - robot simultaneous localization and mapping: a review. *Journal of Field Robotics*.

Schmuck, P., Ziegler, T., Karrer, M., Perraudin, J., & Chli, M. (2021). COVINS: Visual-Inertial SLAM for Centralized Collaboration. *arXiv:2108.05756*.

Turpin, M., N. Michael, and V. Kumar, "Decentralized formation control with variable shapes for aerial robots," in 2012 IEEE international conference on robotics and automation. IEEE, 2012, pp. 23–30.

Wang, C., Zhang, H., Nguyen, T. M., & Xie, L. (2017). Ultra-Wideband Aided Fast Localization and Mapping System. *IEEE/RSJ International Conference on Intelligent Robots & Systems*.

Xu, H, Wang, L., Y. Zhang, K. Qiu and S. Shen, "Decentralized Visual-Inertial-UWB Fusion for Relative State Estimation of Aerial Swarm," 2020 IEEE International Conference on

Robotics and Automation (ICRA), 2020, pp. 8776-8782, doi:
10.1109/ICRA40945.2020.9196944.

Zheng, F., Tang, H., & Liu, Y. H. (2019). Odometry-Vision-
Based Ground Vehicle Motion Estimation With SE(2)-
Constrained SE(3) Poses. IEEE Transactions on Cybernetics,
2652-2663.