

Goal: Enhance vision-based state estimation in multi-UAV systems by exploiting visual feature information during trajectory planning to maintain reliable localization.

Approach: (a) Fisher information-based trajectory evaluation to favor motion that improves expected localization accuracy.

(b) Feature map-based frame alignment to correct relative vision drift and maintain a consistent shared map among agents.

(c) A decentralized planning framework that incorporates both individual feature visibility and covisible landmarks when selecting trajectories.

Motivation

- VIO drift accumulates over time** due to IMU noise and imperfect feature tracking, leading to degraded localization in long-duration operations.
- Feature visibility is highly uneven in realistic environments; entering **feature-sparse regions** can quickly lead to estimator degradation or failure.
- Covisible features** observed across agents provide **strong geometric constraints** that can be exploited to correct relative drift and maintain a consistent shared map.

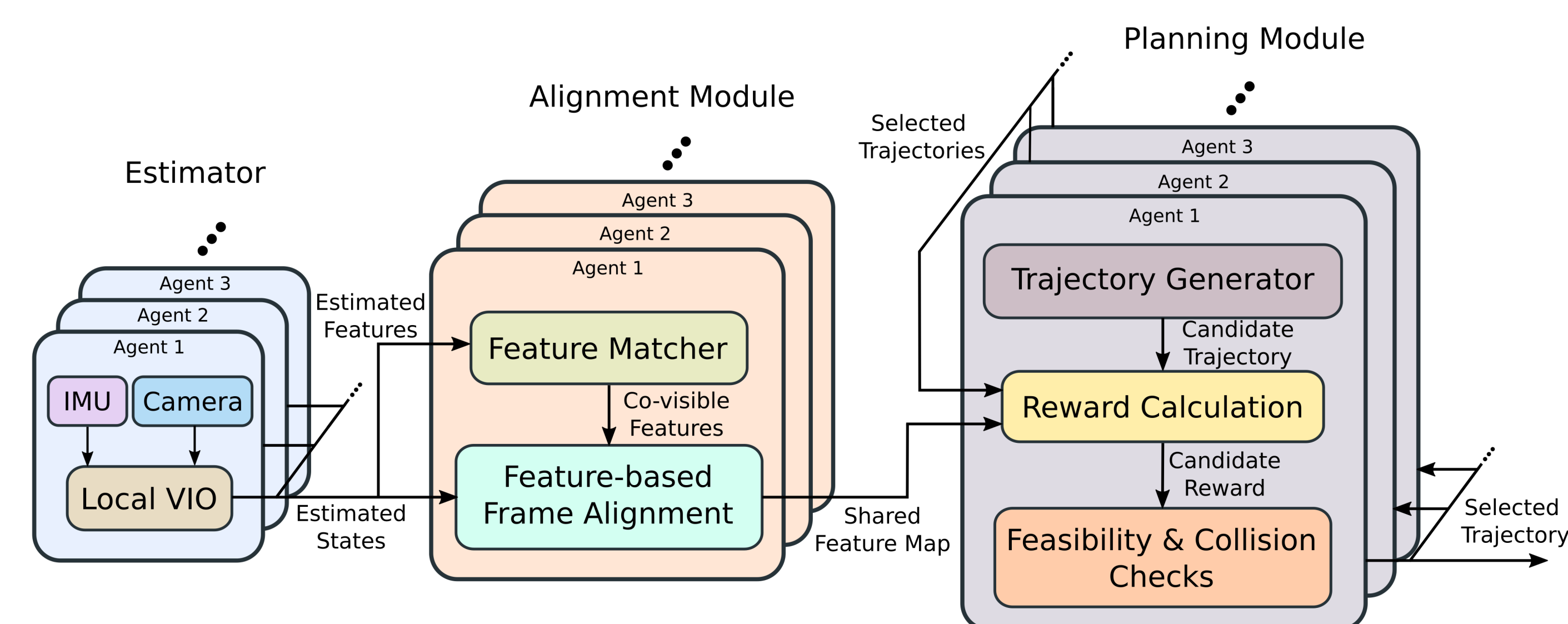
Problem Setup

- Each UAV runs VIO and maintains its own local 3D feature map, which can be shared at communication times.
- Feature-based map alignment fuses shared landmarks to maintain a consistent relative pose across agents.
- The planner selects a trajectory that **advances toward the goal** while **improving expected localization quality**.
- Trajectories are evaluated using a combined reward while ensuring the trajectory is **collision-free** and **dynamically feasible** and is chosen via sampling-based real-time optimization [1].

$$\begin{aligned} \arg \max_{\text{Trajectory } \mathbf{r}_i(t)} \quad & k_{\text{goal}} R_{\text{goal},i} + k_{\text{perc}} R_{\text{perc},i} \\ \text{s.t.} \quad & \mathbf{s}_i(t) \in \mathcal{X}_{\text{free}}, \quad \forall t \in [0, T] \quad \text{Collision-free} \\ & \mathbf{f}_i(t), \omega_i(t) \in \mathcal{U}_{\text{feasible}}, \quad \forall t \in [0, T] \quad \text{Input-feasible} \end{aligned}$$

System Architecture

- Each agent runs a local VIO estimator to generate states and visual features.
- Agents periodically exchange local maps and align covisible features to form a consistent shared feature map.
- The planner samples trajectory candidates, evaluates a combined reward, and selects trajectories based on shared feature map.



Progress Reward

- The goal progress reward encourages the agent to reduce its distance to the target as efficiently as possible.

$$R_{\text{goal},i} = \frac{\|\mathbf{s}_{G,i} - \mathbf{s}(0)\| - \|\mathbf{s}_{G,i} - \mathbf{s}(T)\|}{T}$$

Frame Alignment

- Fuse feature observations from multiple UAVs into a **shared feature map** in the **consensus world frame** to mitigate independent VIO drift.
- Use **pairwise covisible features** to jointly refine agent poses and shared landmarks via nonlinear least-squares optimization.

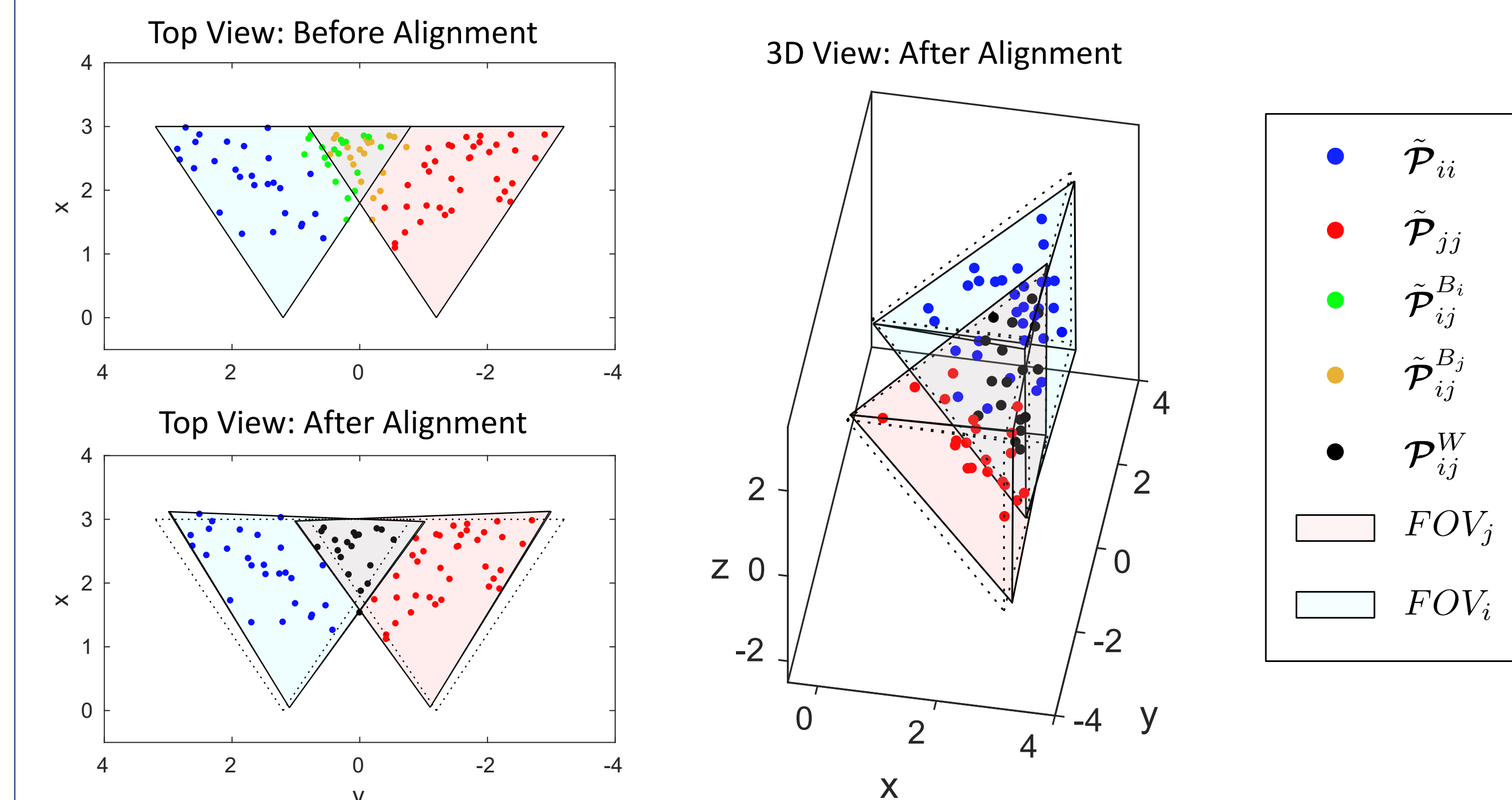
$$\min_{\Theta} \sum_{i \in \mathcal{V}} \|\mathbf{r}_{p,i}\|_{\Sigma_{p,i}^{-1}}^2 + \sum_{(i,j) \in \mathcal{E}} \sum_{k=1}^{m_{ij}} \left(\|\mathbf{r}_{m,i,k}\|_{\Sigma_{m,i}^{-1}}^2 + \|\mathbf{r}_{m,j,k}\|_{\Sigma_{m,j}^{-1}}^2 \right)$$

Decision Variable: $\Theta = \{\xi_i\}_{i \in \mathcal{V}} \cup \{\mathcal{P}_{ij,k}^W\}_{(i,j) \in \mathcal{E}, k=1..m_{ij}}$

Prior Residual: $\mathbf{r}_{p,i} = \xi_i$

Meas. Residual for i : $\mathbf{r}_{m,i,k} = \mathbf{R}^{B_i W} \mathbf{p}_{ij,k}^W + \mathbf{t}^{B_i W} - \tilde{\mathbf{p}}_{ij,k}^{B_i}$

Meas. Residual for j : $\mathbf{r}_{m,j,k} = \mathbf{R}^{B_j W} \mathbf{p}_{ij,k}^W + \mathbf{t}^{B_j W} - \tilde{\mathbf{p}}_{ij,k}^{B_j}$



Results: Frame Alignment

- Two UAVs hovered with injected translational drift (0.5 m) to test the alignment module. The correction reduced inter-agent distance error by **~60%**.

Drift Direction	RMSE Old (m)	RMSE New (m)	% Reduction
X	0.510	0.212	58.3
Y	0.165	0.058	65.0
Z	0.219	0.082	62.7

Information-Based Perception-Aware Reward

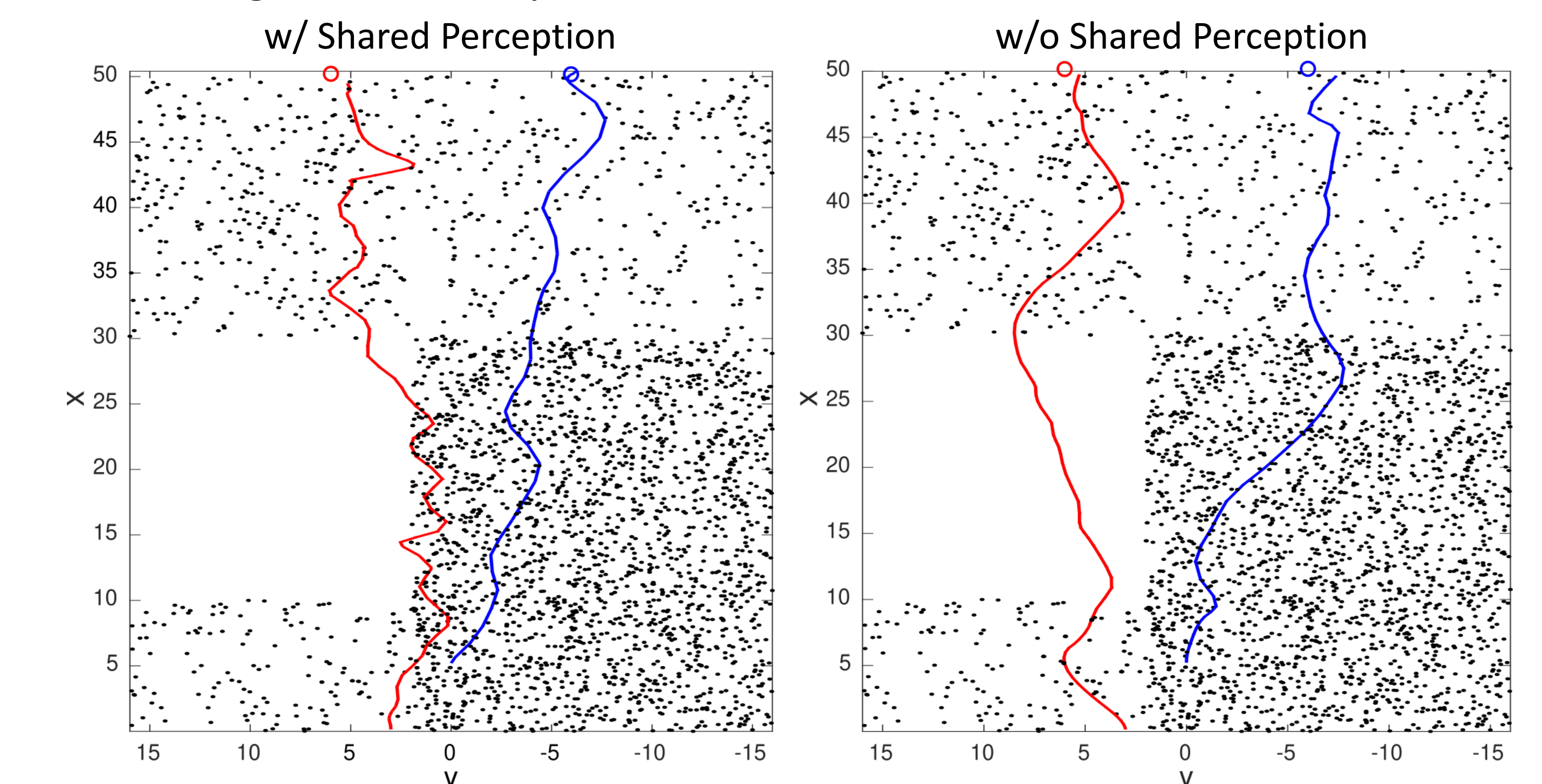
- Use a **Fisher information-based metric** to evaluate the expected information gain along each candidate trajectory.
- Prior information** is constructed following [2], representing the contribution from independent VIO feature tracks.
- Shared measurement information** is derived from the frame alignment least-squares formulation, capturing inter-agent covisibility.
- A **perception-aware reward** is computed using the D-optimality criterion.

$$\Lambda_{\text{meas},i,\tau,k} = J_{\xi_i,k}^P \Sigma_{m,i}^{-1} J_{\xi_i,k}^P \quad J_{\xi_i,k}^P := \frac{\partial \mathbf{p}_{ij,k}^{B_i}}{\partial \xi_i} \bigg|_{\xi_i=0}$$

$$\Lambda_{i,\tau} = \Lambda_{\text{prior},i,\tau} + \Lambda_{\text{meas},i,\tau} \quad R_{\text{perc},i} = \frac{1}{n_{\text{poses}}} \sum_{\tau=1}^{n_{\text{poses}}} \log \det(\Lambda_{i,\tau})$$

Results: Receding-Horizon Planning

- Planning with shared perception helps agents **avoid feature-poor regions** that would degrade odometry.



- Simulations of two UAVs planning in three environments compared perception-aware and baseline methods. The proposed planner consistently **increased visible and shared features**, enhancing observability while maintaining **comparable goal accuracy**.

Case	Method	# Visible	# Covisible	Error A	Error B
1	perception-agnostic	149	18	0.52	0.56
	perception-aware	169	30	0.80	0.56
2	perception-agnostic	132	0	0.23	0.24
	perception-aware	141	0	0.39	0.50
3	perception-agnostic	43	3	1.07	1.06
	perception-aware	72	8	0.96	1.01

References

- N. Bucki, J. Lee, and M. W. Mueller, "Rectangular pyramid partitioning using integrated depth sensors (RAPIDS): A fast planner for multicopter navigation," IEEE Robotics and Automation Letters, vol. 5, no. 3, pp. 4626–4633, 2020.
- X. Wu, S. Chen, K. Sreenath, and M. W. Mueller, "Perception-aware receding horizon trajectory planning for multicopters with visual-inertial odometry," IEEE Access, Vol 10, pp. 87911-87922, 2022.