FEJ2: A Consistent Visual-Inertial State Estimator Design

Chuchu Chen, Yulin Yang, Patrick Geneva, and Guoquan Huang

University of Delaware

ICRA 2022 IEEE International Conference on Robotics and Automation

Motivation & Contribution

- Filter-based visual-inertial navigation systems (VINS) fuse IMU readings and camera measurements
  - Perform linearization at **current** state estimates
  - **Inconsistent:** Overconfident covariance due to information gains along unobservable directions
  - 4 d.o.f **unobservable** directions (yaw + position)
- First-Estimates Jacobian (FEJ) [1]
  - Enforce ideal unobservable directions and improve estimator consistency
  - Fix Jacobian evaluations at **first** state estimates but possibly introduce **unmodelled errors**
- We propose **FEJ2**
  - Address **unmodelled errors** introduced by FEJ
  - **Improve** performance for high noise cases

Standard VINS

- Linearize the nonlinear measurement function at the **current** state estimate $\hat{x}$ for EKF update:

$$z = h(x) + n \simeq h(\hat{x}) + \hat{H}\tilde{x} + n$$

- Observability analysis shows using $\hat{x}$ causes the filter to gain **extra information inconsistently**

First-Estimates Jacobian (FEJ)-VINS

- Evaluate the measurement Jacobian at the **first** state estimate x for EKF update:

$$z = h(x) + n \simeq h(\hat{x}) + H\tilde{x} + n$$

- Enforce **correct** unobservable directions and improve system **consistency**, but introduces unmodelled errors:

$$\begin{aligned} z &\simeq h(x) + H(x \quad x) + n \\ &= h(x) + H(\hat{x} \quad x) + H(x \quad \hat{x}) + n \\ &\simeq h(\hat{x}) + H\tilde{x} + n \end{aligned}$$

- Thus, FEJ assumes the following:

$$\boxed{h(\hat{x}) \simeq h(x) + H(\hat{x} \quad x)}$$

FEJ2-VINS

1. Linearize at the **current** state estimates $\hat{x}$ for the smallest linearization errors
2. Use the **first-estimates Jacobian** H to avoid extra information gain along unobservable directions
3. Derive a more **accurate** linear model:

$$\begin{aligned} z &\simeq h(\hat{x}) + \hat{H}\tilde{x} + n \\ &= h(\hat{x}) + (H + \hat{H} \quad H)\tilde{x} + n \\ r = z \quad h(\hat{x}) &= H\tilde{x} + \quad H\tilde{x} + n \end{aligned} \qquad (1)$$

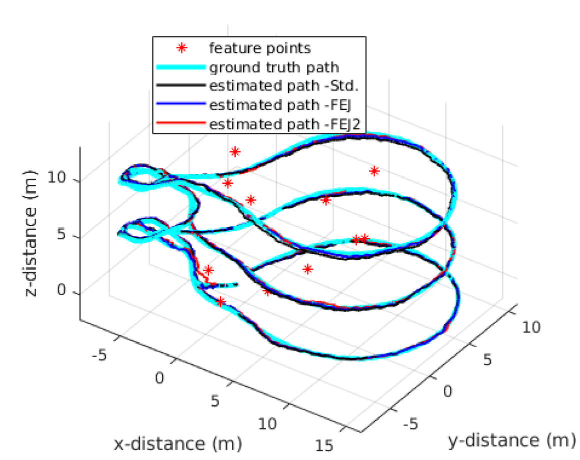
4. $\quad H = \hat{H} \quad H$ captures the errors of using the first-estimates Jacobians for update which is **ignored** by FEJ
5. **Marginalize** $\quad H$ by projecting Eq. (1) onto its left nullspace U to **compensate the unmodelled errors** of FEJ

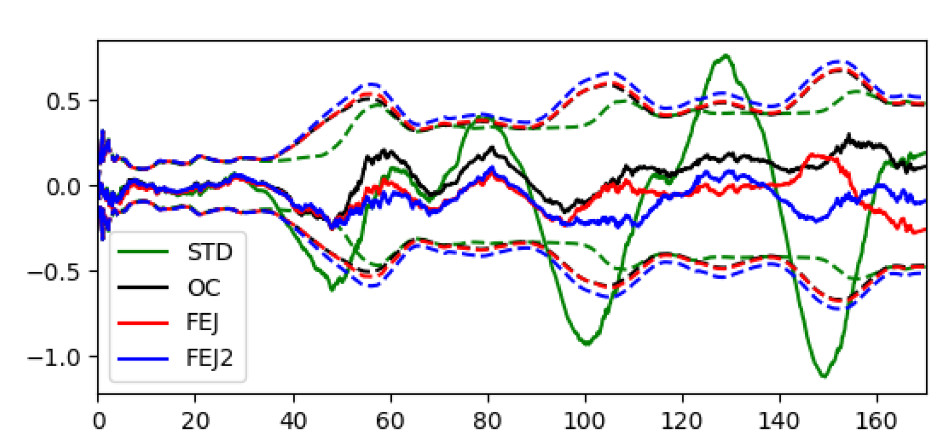
$$\begin{aligned} U^{\top} r &= \quad U^{\top} H\tilde{x} + \cancel{U^{\top} \quad H\tilde{x}} + \quad U^{\top} n \\ r \quad &= H \quad \tilde{x} + n \end{aligned}$$

- **FEJ2 Properties:**
  - FEJ2 has an unobservable subspace of **correct** dimensions and structure, and shares the **same** (initial) nullspace of the observability matrix as the FEJ and Observability-constrained (OC)-VINS [2].
  - FEJ2 has **larger** covariance estimates (better consistency) than FEJ

SLAM Simulations

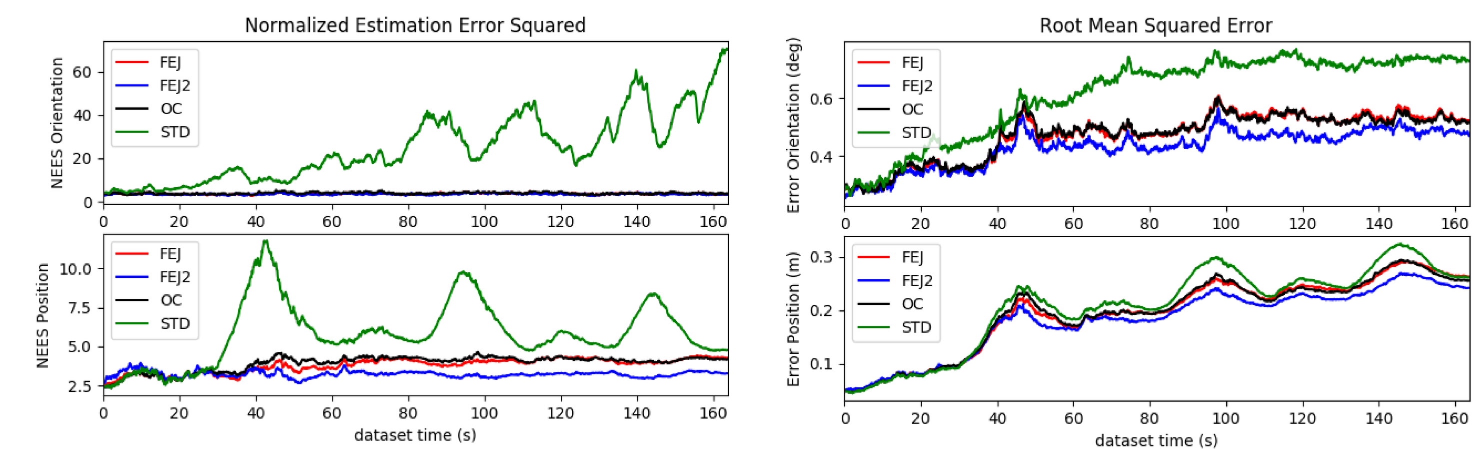
- The standard (STD) estimator is **inconsistent** with estimation errors frequently out of the 3 bounds
- The errors for OC, FEJ and FEJ2 are **consistently** within their 3 bounds
- The 3 bounds of FEJ2 are **slightly broader**





MSCKF-VIO Simulations

- Test with both monocular and stereo measurements under different noise levels within OpenVINS [3]
- FEJ2 achieves the **smallest estimate errors** and the **most ideal NEES values**, especially when measurement noise is large



Noise (pixel)	Est.	RMSE Ori. (deg) mono / stereo	RMSE Pos. (m) mono / stereo	NEES Ori. mono / stereo	NEES Pos. mono / stereo
1	STD	0.412 / 0.344	0.130 / 0.109	23.874 / 15.447	4.911 / 4.874
	OC	0.242 / 0.257	0.119 / 0.100	3.290 / 3.599	3.540 / 3.416
	FEJ	0.242 / 0.256	0.120 / 0.100	3.284 / 3.438	3.617 / 3.322
	FEJ2	**0.238 / 0.238**	**0.118 / 0.095**	**3.150 / 3.324**	**3.443 / 2.965**
3	STD	2.139 / 0.888	0.402 / 0.310	407.221 / 33.852	13.212 / 7.235
	OC	0.716 / 0.723	0.301 / 0.300	3.964 / 4.395	5.051 / 4.839
	FEJ	0.861 / 0.704	0.289 / 0.298	4.965 / 4.163	4.763 / 4.656
	FEJ2	**0.650 / 0.663**	**0.264 / 0.277**	**3.198 / 3.790**	**3.581 / 3.636**

Real-World Experiments

- Test on the Euroc Mav dataset and TUM-VI dataset with both monocular and stereo configurations
- FEJ2 achieves **better accuracy** on average, especially in monocular camera scenarios
- Report absolute trajectory error (ATE) for each estimator in units of degree/meters averaged over 10 runs

	Est.	V1_01_easy	V1_02_med.	V1_03_dif	V2_01_easy	V2_02_med.	V2_03_dif.
mono	STD	0.956 / 0.076	1.783 / 0.080	2.638 / 0.074	0.951 / 0.098	1.856 / 0.085	**1.415** / 0.154
	OC	0.554 / 0.077	0.615 / 0.071	2.933 / 0.068	0.880 / 0.093	1.595 / 0.077	1.835 / 0.188
	FEJ	0.872 / 0.056	0.574 / **0.052**	**2.079** / 0.096	0.928 / **0.092**	1.599 / 0.074	1.874 / 0.168
	FEJ2	**0.679** / **0.053**	**0.564** / 0.059	2.346 / **0.061**	**0.791** / 0.101	**1.233** / **0.047**	1.808 / **0.146**
stereo	STD	0.792 / 0.061	1.958 / 0.059	2.551 / 0.053	1.078 / 0.055	1.693 / 0.064	2.337 / 0.077
	OC	0.615 / 0.071	1.772 / 0.046	**2.468** / **0.045**	1.098 / 0.059	1.231 / 0.051	**1.052** / 0.061
	FEJ	**0.547** / **0.052**	**1.702** / 0.079	2.498 / **0.045**	1.172 / 0.058	1.268 / 0.049	1.118 / **0.058**
	FEJ2	0.564 / 0.059	1.770 / **0.045**	2.503 / 0.047	**0.975** / **0.053**	**1.202** / **0.047**	1.101 / 0.062

[1] Huang GP, Mourikis AI and Roumeliotis SI. Observability-based Rules for Designing Consistent EKF SLAM Estimators. The International Journal of Robotics Research (IJRR), 2010.
[2] Hesch JA, Kottas DG, Bowman SL, Roumeliotis SI. Camera-IMU-based localization: Observability analysis and consistency improvement. The International Journal of Robotics Research (IJRR), 2014.
[3] P. Geneva, K. Eckenhoff, W. Lee, Y. Yang and G. Huang, "OpenVINS: A Research Platform for Visual-Inertial Estimation," 2020 IEEE International Conference on Robotics and Automation (ICRA), 2020.
https://github.com/rpng/open_vins