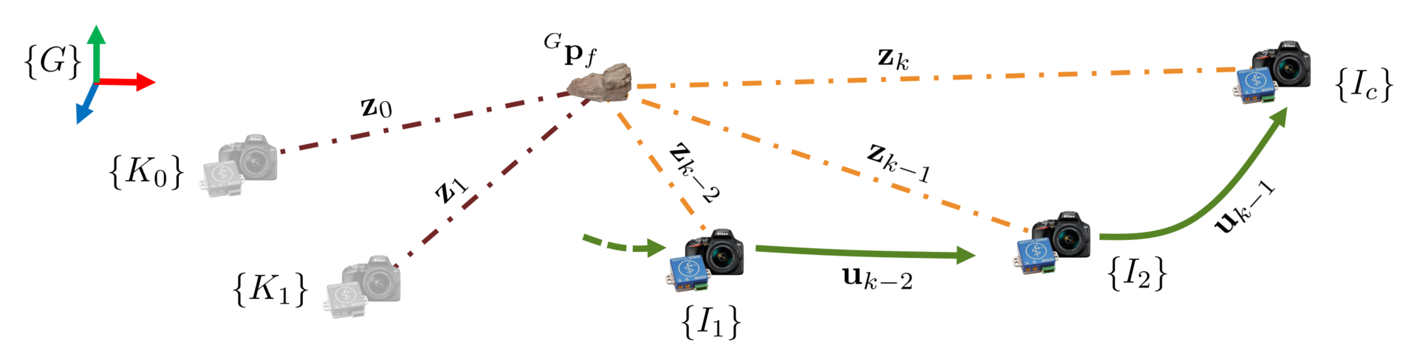
Map-based Visual-Inertial Localization: A Numerical Study

Patrick Geneva and Guoquan Huang

Motivation & Contribution

- Leverage prior map information to improve visual-inertial navigation systems (VINS)
  - Prior point map with 2D-to-3D meas.
  - Prior keyframe map with 2D-to-2D meas.
- Summary of different filtering techniques for incorporating loop-closures
  - Extended Kalman filters (EKF)
  - Schmidt-Kalman filters (SKF) [1]
  - Measurement inflation models (INF)
- Investigate the accuracy, consistency, computational complexity, and memory

Map-based EKF-SLAM



Goal: Leverage a prior map to constrain "active" sliding window filter in an efficient manner.

$$\mathbf{x}_k = \begin{bmatrix}\mathbf{x}_A & \mathbf{x}_M & \mathbf{x}_K\end{bmatrix}, \quad \mathbf{x}_A = \begin{bmatrix}\mathbf{x}_{I_k} & \mathbf{x}_C & \mathbf{x}_L\end{bmatrix}$$

We have the following point or keyframe map:

$$\mathbf{x}_M = \begin{bmatrix}{}^G\mathbf{p}_{f_1} & \cdots & {}^G\mathbf{p}_{f_m}\end{bmatrix}, \quad \mathbf{x}_K = \begin{bmatrix}\mathbf{x}_{T_1} & \cdots & \mathbf{x}_{T_n}\end{bmatrix}$$

State $\mathbf{x}_{T_i} = [{}^{I_i}_{G}\bar{q} \;\; {}^G\mathbf{p}_{I_i}]$ and environmental feature related through bearing projection:

$$\mathbf{z}_k = \mathbf{h}(\mathbf{x}_{T_k}, {}^G\mathbf{p}_f) + \mathbf{n}_k =: \mathbf{\Lambda}({}^{C_k}\mathbf{p}_f) + \mathbf{n}_k$$

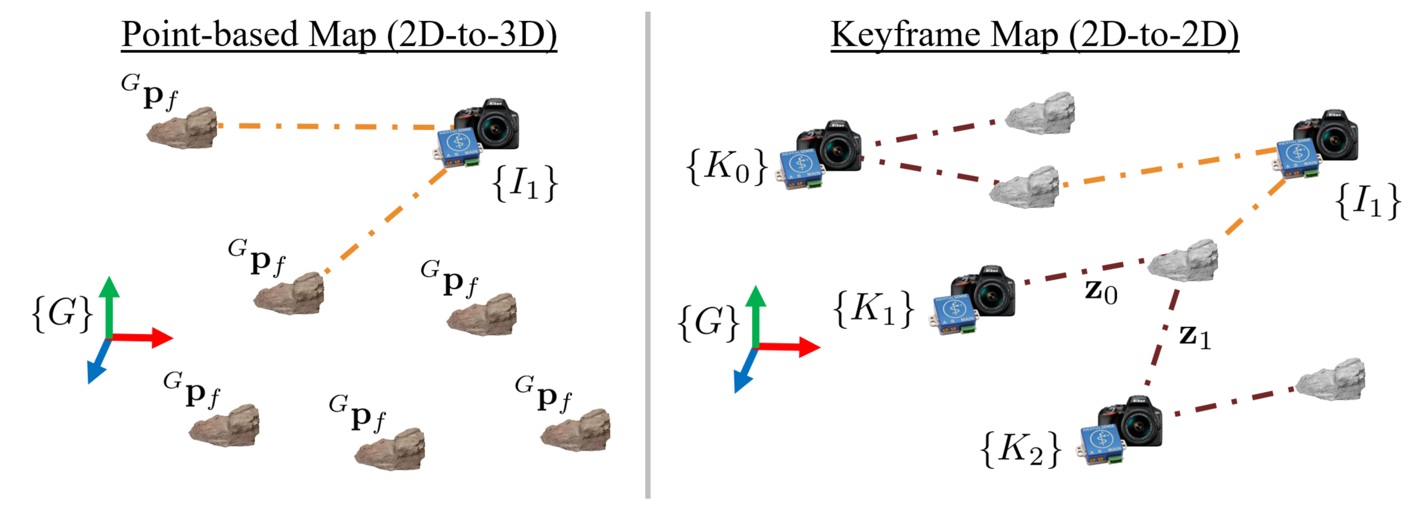
$${}^{C_k}\mathbf{p}_f = {}^C_I\mathbf{R}\,{}^{I_k}_G\mathbf{R}({}^G\mathbf{p}_f - {}^G\mathbf{p}_{I_k}) + {}^C\mathbf{p}_I$$

We can linearize this function to get the following:

$$\mathbf{r}_k = \mathbf{z}_k - \mathbf{h}(\hat{\mathbf{x}}_{T_k}, {}^G\hat{\mathbf{p}}_f) \simeq \mathbf{H}_{T_k}\tilde{\mathbf{x}}_{T_k} + \mathbf{H}_{f_k}{}^G\tilde{\mathbf{p}}_f + \mathbf{n}_k$$

Thus the update is a function of the **observation pose** and **feature position**, either of which can be in our prior map. The measurement noise for an observation is $\mathbf{n}_k \sim \mathcal{N}(\mathbf{0}, \sigma_{pix}^2\mathbf{I})$.

Prior Map Methodologies



- Point-based map (PTS) uses 2D-to-3D measurement of **3d feature positions** [2]
- Keyframe-based (KFS) map uses 2D-to-2D map **bearing observation** and keyframe 6dof pose [3]

Efficient Prior Map Localization

- Extended Kalman filter (EKF), updates map, tracks full correlations, **expensive** computation
- Schmidt-Kalman filter (SKF), map states not updated, correlations are, **proven consistency**

$$\hat{\mathbf{x}}_{A_k}^{\oplus} = \hat{\mathbf{x}}_{A_k} + \mathbf{K}_{A_k}\mathbf{r}, \quad \hat{\mathbf{x}}_{S_k}^{\oplus} = \hat{\mathbf{x}}_{S_k}$$

$$\mathbf{P}_k^{\oplus} = \mathbf{P}_k - \begin{bmatrix} \mathbf{K}_{A_k}\mathbf{S}_k\mathbf{K}_{A_k}^\top & \mathbf{K}_{A_k}\mathbf{H}_k\begin{bmatrix}\mathbf{P}_{AS_k}\\ \mathbf{P}_{SS_k}\end{bmatrix} \\ \begin{bmatrix}\mathbf{P}_{AS_k}\\ \mathbf{P}_{SS_k}\end{bmatrix}^\top \mathbf{H}_k{}^\top\mathbf{K}_{A_k}^\top & 0 \end{bmatrix}$$

- Noise inflation methods, treat feature position or pose as **true**, compensate through inflation
  - Measurement $\mathbf{R} = (\gamma\sigma_{pix})^2\mathbf{I}$
  - Marginal Covariance
    $$\mathbf{R} = \mu\mathbf{H}'_{T_k}\mathbf{P}_{TT_k}\mathbf{H}'^\top_{T_k} + \sigma_{pix}^2\mathbf{I}$$
  - Alpha-Beta Covariance
    $$\mathbf{R} = \alpha\mathbf{H}'_{T_k}\mathbf{P}_{TT_k}\mathbf{H}'^\top_{T_k} + \beta\mathbf{H}'_T\mathbf{P}\mathbf{H}'^\top_T + \sigma_{pix}^2\mathbf{I}$$

Algo.	Cxty PTS	Cxty KFS	Mem. PTS	Mem. KFS
EKF	$O(9m^2)$	$O(36n^2)$	$O(9m^2)$	$O(36n^2)$
SKF	$\mathbf{O(3m)}$	$O(6n)$	$O(9m^2)$	$O(36n^2)$
Inf. γ	$O(1)$	$O(1)$	$O(0)$	$O(0)$
Inf. μ	$\mathbf{O(1)}$	$O(1)$	$O(9m)$	$O(36n)$
Inf. α,β	$O(1)$	$O(1)$	$O(9m)$	$O(36n)$

Simulation – Prior Map Accuracy

- Prior map points generated based on trajectory with keyframes generated from prior distribution [4]
- EKF and SKF can **gain accuracy** improvements over VIO at all noise levels
- Even **significantly noisy** prior maps can still provide information

	Prior	Algo.	ATE (deg / m)	NEES (3)
	-	VIO	**2.603 / 0.271**	3.524 / 1.591
2D-to-2D	1.0°, 6cm	EKF	0.442 / 0.105	3.236 / 3.698
		SKF	0.518 / 0.130	2.806 / 3.466
	3.0°, 12cm	EKF	**0.629 / 0.127**	4.353 / 5.335
		SKF	0.941 / 0.167	3.009 / 3.585
2D-to-3D	6cm	EKF	0.068 / 0.014	8.224 / 9.292
		SKF	0.087 / 0.036	2.863 / 3.210
	12cm	EKF	**0.079 / 0.015**	9.321 / 9.472
		SKF	0.122 / 0.065	2.761 / 3.175

Simulation – Method Comparisons

- All methods are able to **outperform** standard visual-inertial odometry
- SKF and inflation-based methods are most **efficient**
- After tuning, inflation methods are consistent, with low computational cost, while **SKF-based guarantees consistency** with a fraction of the cost of the EKF-based

	Algo.	40m	80m	120m	160m	200m	240m	NEES (ori / pos)	Time (ms)
	VIO	0.373 / 0.088	0.536 / 0.119	0.636 / 0.141	**0.717 / 0.163**	**0.811 / 0.175**	0.888 / 0.187	3.228 / 3.796	0.8 ± 0.3
2D-to-2D	EKF	0.225 / 0.091	0.323 / 0.111	0.372 / 0.120	**0.402 / 0.121**	0.424 / 0.122	0.394 / 0.125	3.298 / 4.311	3.6 ± 1.8
	SKF	0.260 / 0.097	0.339 / 0.129	0.415 / 0.146	**0.448 / 0.155**	0.492 / 0.167	0.542 / 0.171	3.074 / 3.596	1.4 ± 0.7
	Inf. Meas.	0.276 / 0.099	0.353 / 0.134	0.449 / 0.152	**0.518 / 0.163**	0.531 / 0.173	0.562 / 0.180	3.016 / 3.647	0.9 ± 0.3
	Inf. Marg.	0.265 / 0.091	0.350 / 0.122	0.447 / 0.142	**0.520 / 0.156**	0.560 / 0.169	0.613 / 0.175	2.795 / 2.784	0.9 ± 0.3
	Inf. $\alpha\beta$	0.269 / 0.091	0.353 / 0.122	0.456 / 0.142	**0.546 / 0.156**	0.599 / 0.168	0.656 / 0.173	2.781 / 2.689	0.9 ± 0.3
2D-to-3D	EKF	0.041 / 0.009	0.041 / 0.009	0.041 / 0.009	0.041 / 0.009	**0.041 / 0.009**	0.041 / 0.009	9.612 / 7.792	5.8 ± 1.1
	SKF	0.090 / 0.040	0.092 / 0.038	0.091 / 0.040	0.090 / 0.038	**0.092 / 0.039**	0.091 / 0.039	3.051 / 2.963	1.4 ± 0.2
	Inf. Meas.	0.125 / 0.068	0.139 / 0.065	0.141 / 0.067	0.141 / 0.064	**0.142 / 0.066**	0.136 / 0.065	3.663 / 3.528	0.6 ± 0.1
	Inf. Marg.	0.102 / 0.046	0.103 / 0.045	0.102 / 0.046	0.098 / 0.044	**0.103 / 0.046**	0.100 / 0.045	3.201 / 2.546	0.6 ± 0.1
	Inf. $\alpha\beta$	0.102 / 0.047	0.103 / 0.046	0.102 / 0.047	0.098 / 0.045	**0.103 / 0.046**	0.100 / 0.046	3.126 / 2.437	0.6 ± 0.1

Summary and Discussion

- Prior maps can be leveraged at even high noise levels to improve accuracy
- Point-based **2D-to-3D** provide the **most accuracy** improvements
- Keyframes **2D-to-2D reduce computation** for large number of features
- SKF is computationally **efficient** and **consistent**
- Inflation methods are relatively invariant to their chosen parameters and can handle **large maps**

[1] Schmidt, Stanley "Application of state-space methods to navigation problems." Advances in control systems. Vol. 3. Elsevier, 1966.
[2] DuToit, Ryan C., et al. "Consistent map-based 3D localization on mobile devices." ICRA. IEEE, 2017.
[3] Geneva, Patrick, et al. "A linear-complexity EKF for visual-inertial navigation with loop closures." ICRA. IEEE, 2019.
[4] Geneva, Patrick, et al. "OpenVINS: A research platform for visual-inertial estimation." ICRA. IEEE, 2020.