Monocular Visual-Inertial Odometry with Planar Regularities

**Chuchu Chen*, Patrick Geneva*,
Yuxiang Peng, Woosik Lee, and Guoquan Huang**

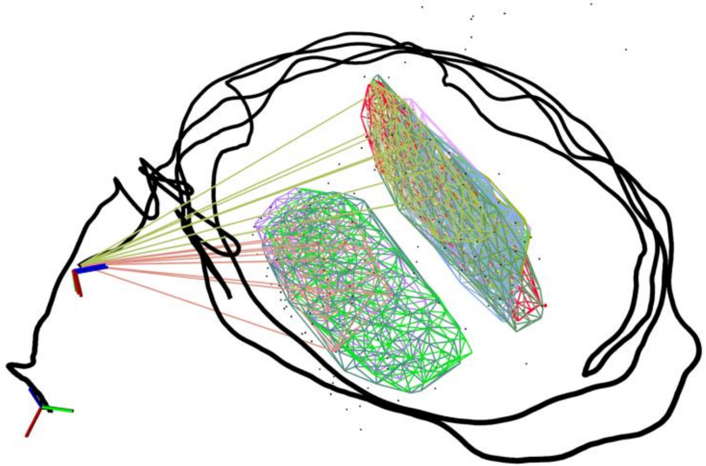
Robot Perception and Navigation Group (RPNG)
University of Delaware, USA

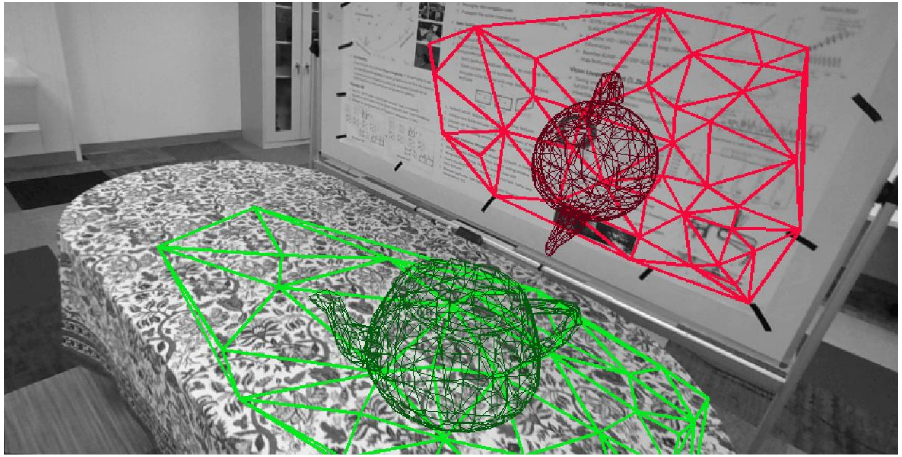
ICRA
LONDON · 2023

RPNG

Introduction

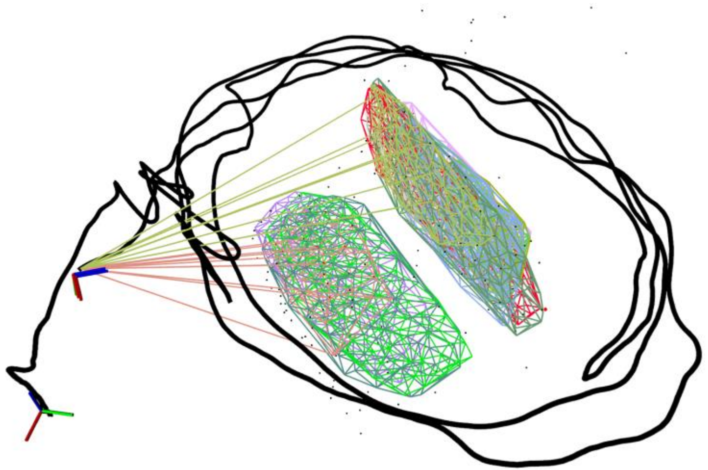
- Man-made environments provide **rich structural** information (e.g. planes...)
- Estimation of **planes** enables scene understanding (e.g. AR / VR)
- Constraining 3D features to planes through point-on-plane **regularization** can improve efficiency (reduce state)





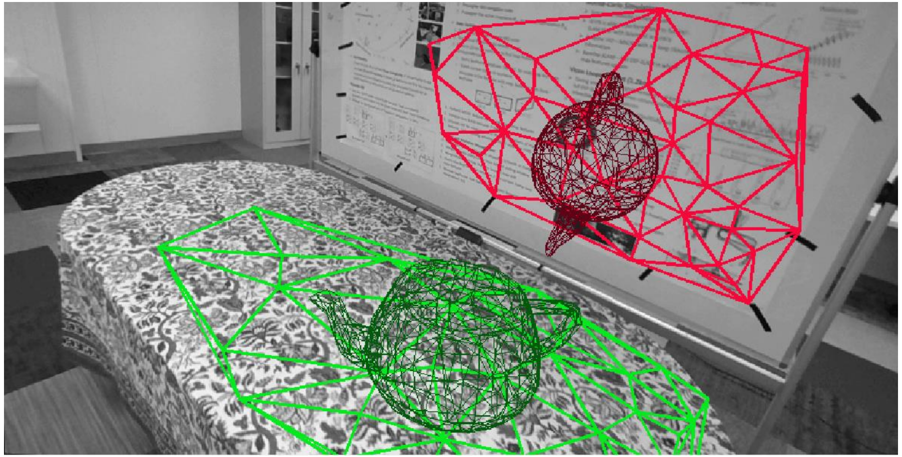
Introduction

- Man-made environments provide **rich structural** information (e.g. planes...)
- Estimation of **planes** enables scene understanding (e.g. AR / VR)
- Constraining 3D features to planes through point-on-plane **regularization** can improve efficiency (reduce state)



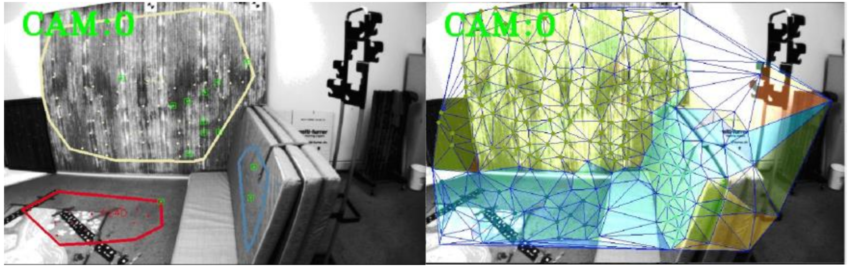
Key Contributions:

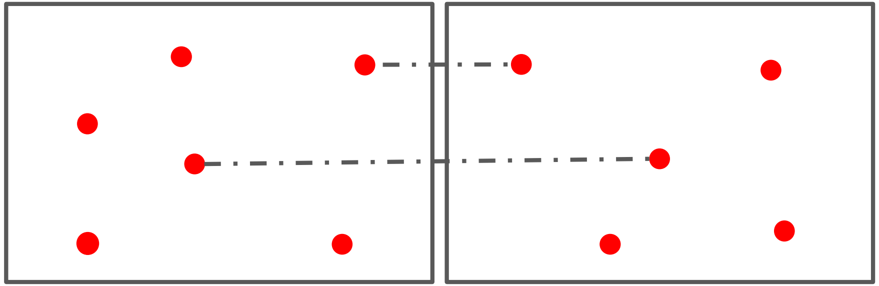
- Novel plane detection and tracking with only a **monocular** camera
- **Efficient** filter-based VIO with planar regularities
- **Open sourced** code and dataset



Monocular Plane Feature Detection and Tracking

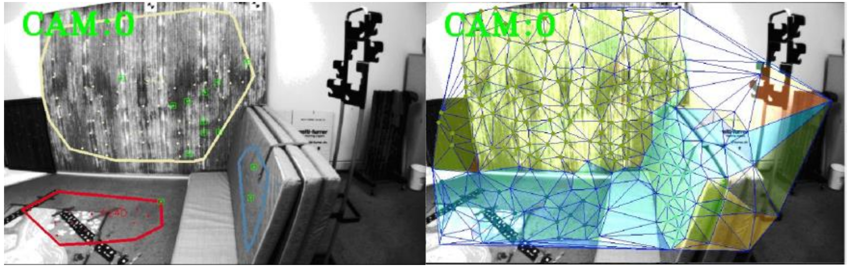
- **Sparse Point Features:**
  - FAST detection
  - KLT provides frame-to-frame tracking

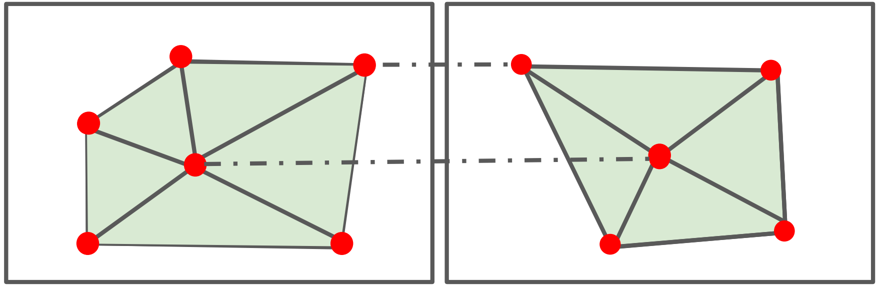




Monocular Plane Feature Detection and Tracking

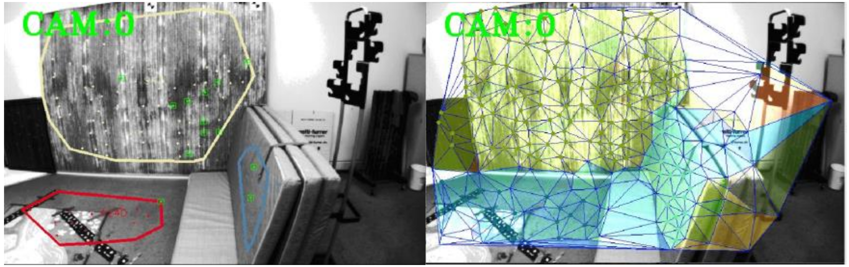
- **Sparse Point Features:**
  - FAST detection
  - KLT provides frame-to-frame tracking
- **Point Feature Meshing**
  - 3D point feature recovery
  - 2D mesh (Delaunay triangulation)
  - Compute normals of triangles

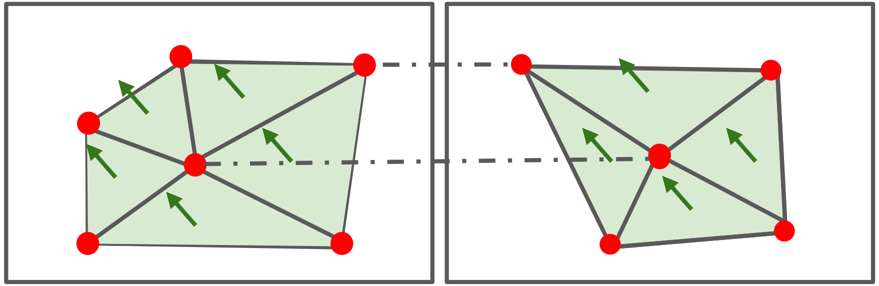




Monocular Plane Feature Detection and Tracking

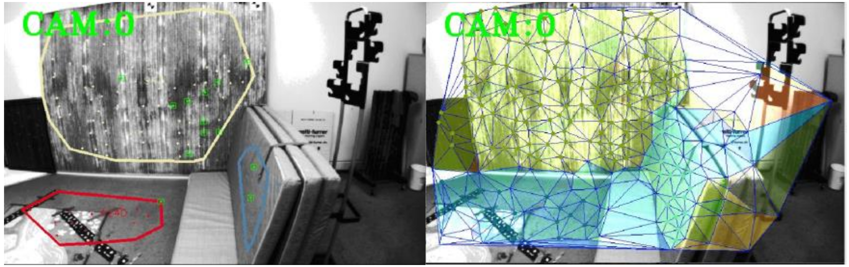
- **Sparse Point Features:**
  - FAST detection
  - KLT provides frame-to-frame tracking
- **Point Feature Meshing**
  - 3D point feature recovery
  - 2D mesh (Delaunay triangulation)
  - Compute normals of triangles

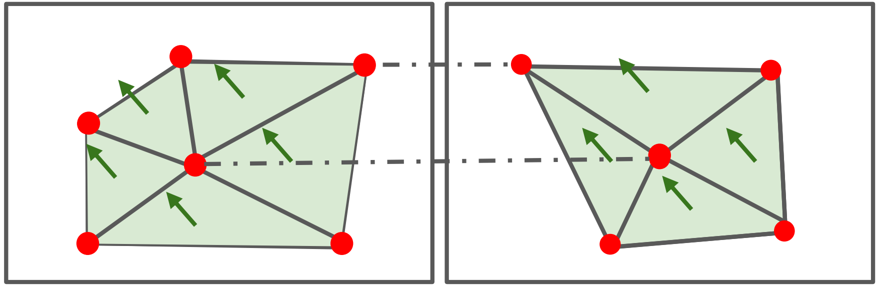




Monocular Plane Feature Detection and Tracking

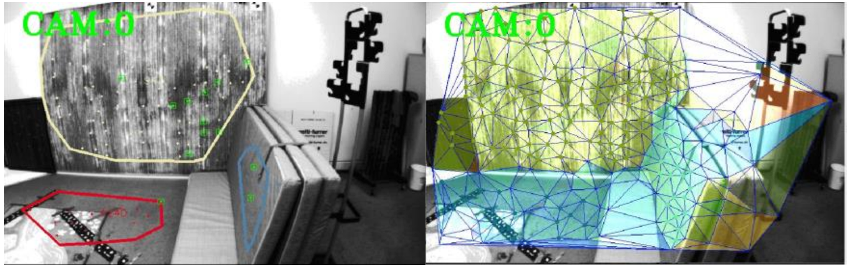
- **Sparse Point Features:**
  - FAST detection
  - KLT provides frame-to-frame tracking
- **Point Feature Meshing**
  - 3D point feature recovery
  - 2D mesh (Delaunay triangulation)
  - Compute normals of triangles
- **Vertex Normals:**
  - Use neighboring triangles normals to compute avg. vertex normals
  - Remove invalidate vertices with high variance (points on the edge)

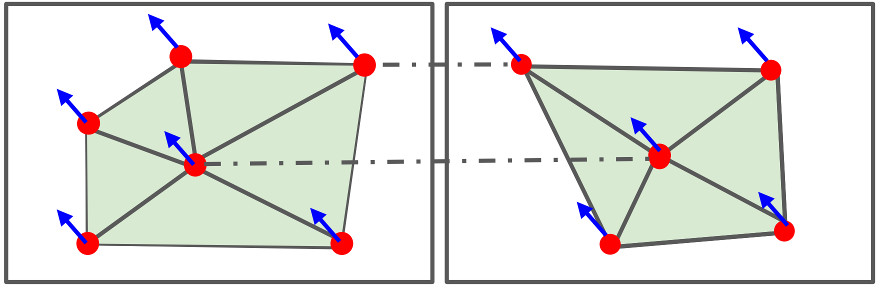




Monocular Plane Feature Detection and Tracking

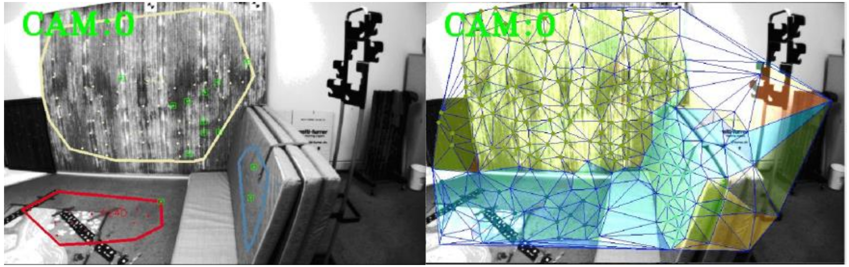
- **Sparse Point Features:**
  - FAST detection
  - KLT provides frame-to-frame tracking
- **Point Feature Meshing**
  - 3D point feature recovery
  - 2D mesh (Delaunay triangulation)
  - Compute normals of triangles
- **Vertex Normals:**
  - Use neighboring triangles normals to compute avg. vertex normals
  - Remove invalidate vertices with high variance (points on the edge)

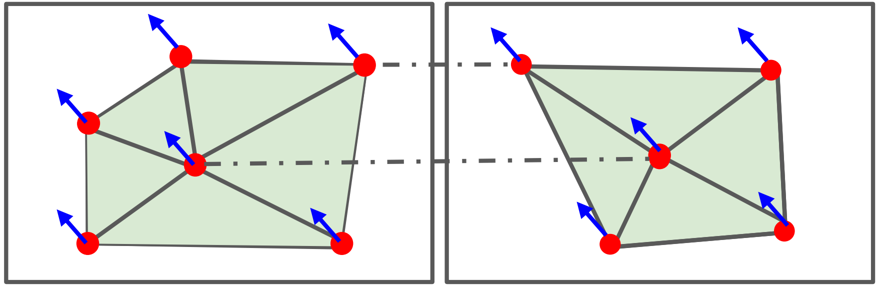




Monocular Plane Feature Detection and Tracking

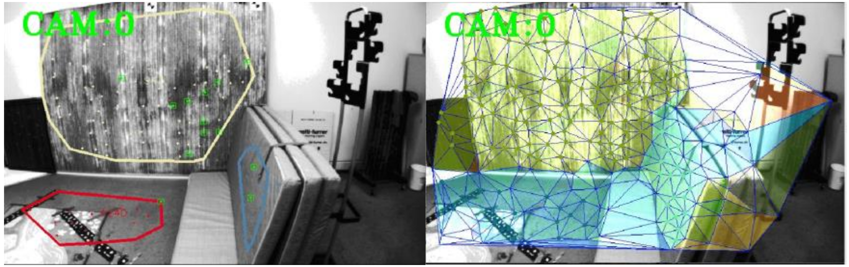
- **Sparse Point Features:**
  - FAST detection
  - KLT provides frame-to-frame tracking
- **Point Feature Meshing**
  - 3D point feature recovery
  - 2D mesh (Delaunay triangulation)
  - Compute normals of triangles
- **Vertex Normals:**
  - Use neighboring triangles normals to compute avg. vertex normals
  - Remove invalidate vertices with high variance (points on the edge)
- **Vertex Matching Heuristics**
  - Pairwise comparison with neighbors
  - Normal difference, point-to-plane distance, statistical filter to remove outlier

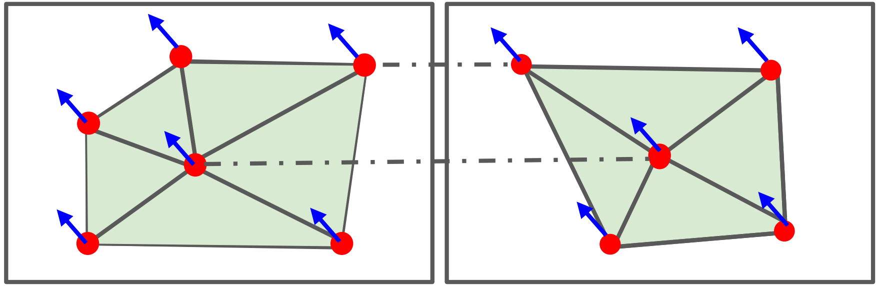




Monocular Plane Feature Detection and Tracking

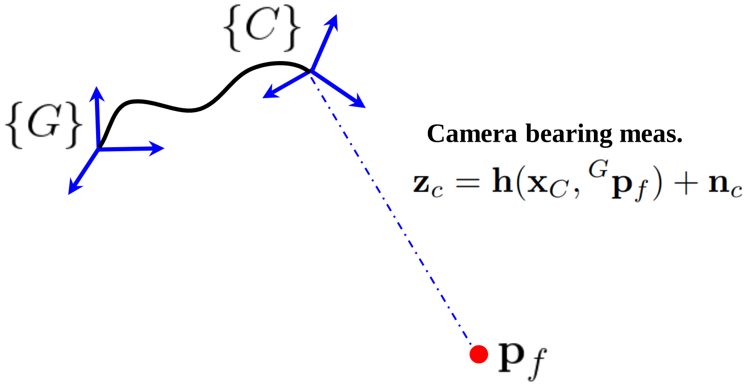
- **Sparse Point Features:**
  - FAST detection
  - KLT provides frame-to-frame tracking
- **Point Feature Meshing**
  - 3D point feature recovery
  - 2D mesh (Delaunay triangulation)
  - Compute normals of triangles
- **Vertex Normals:**
  - Use neighboring triangles normals to compute avg. vertex normals
  - Remove invalidate vertices with high variance (points on the edge)
- **Vertex Matching Heuristics**
  - Pairwise comparison with neighbors
  - Normal difference, point-to-plane distance, statistical filter to remove outlier





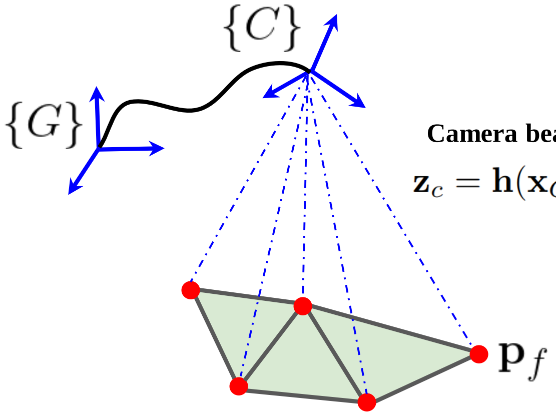
Shown efficient and robust from real-world experiments
(e.g. ~*3-4 ms* on EuRoC Mav)

Planar Regularities



Point: ${}^G\mathbf{p}_f$

Planar Regularities

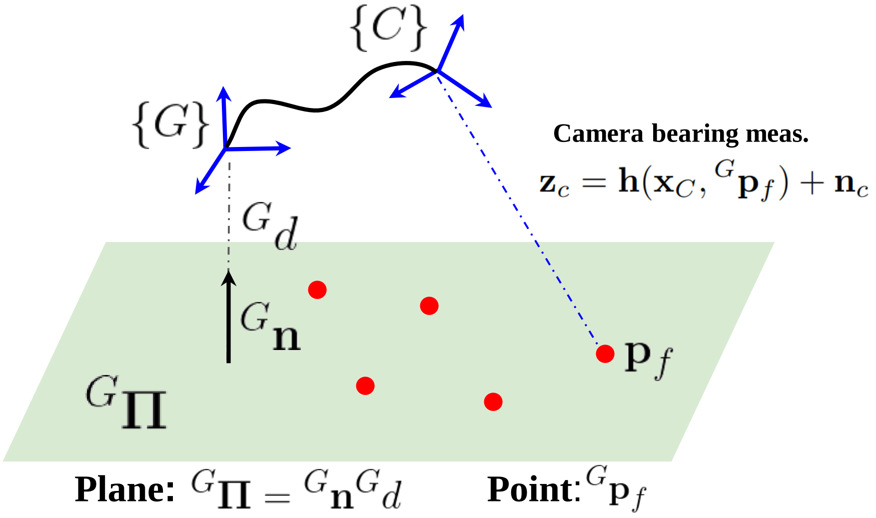


Camera bearing meas.

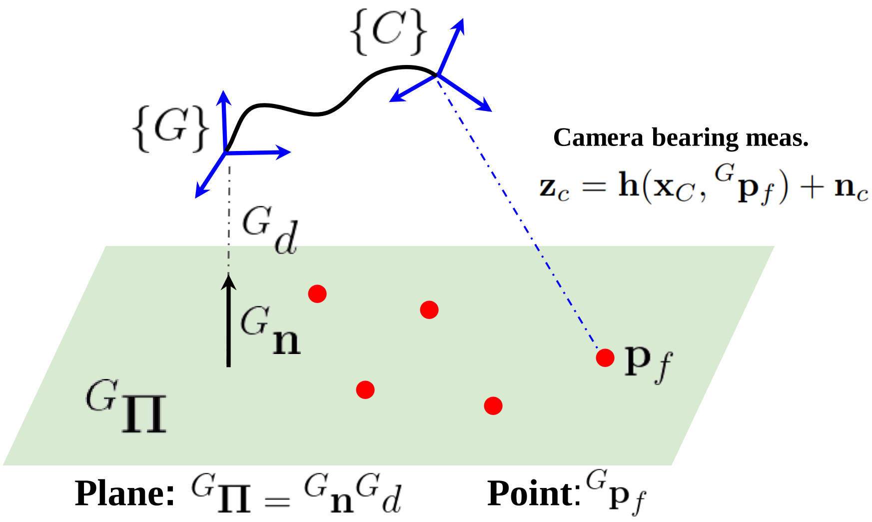
$$\mathbf{z}_c = \mathbf{h}(\mathbf{x}_C, {}^G\mathbf{p}_f) + \mathbf{n}_c$$

Point: ${}^G\mathbf{p}_f$

Planar Regularities



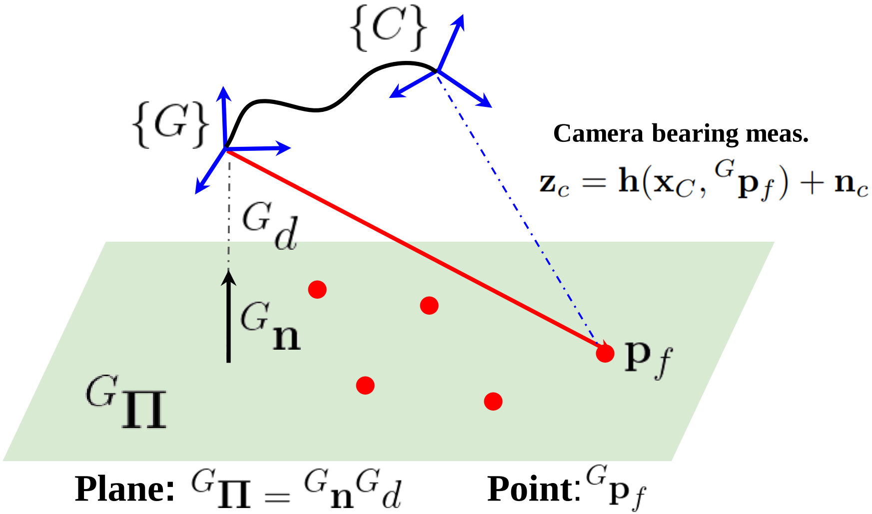
Planar Regularities



Point-on-plane regularity

$$z_d = ({}^G\mathbf{p}_f^\top \, {}^G\mathbf{n} - {}^G d) + n_d$$

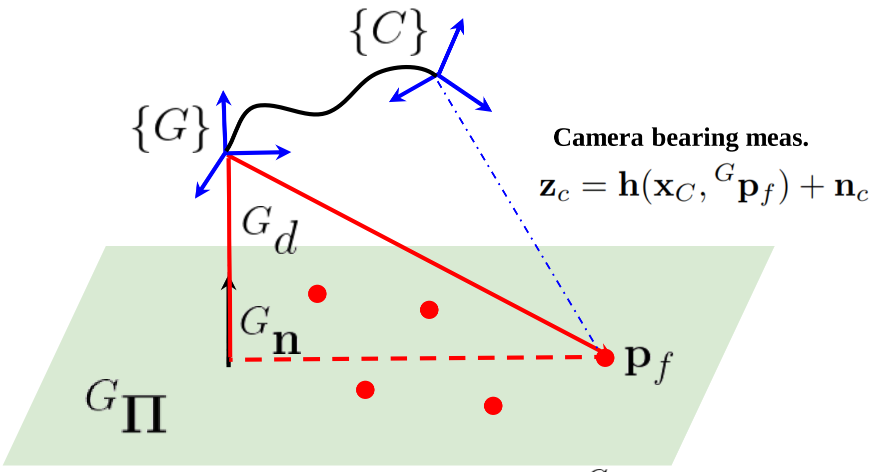
Planar Regularities



Point-on-plane regularity

$$z_d = ({}^G\mathbf{p}_f^\top {}^G\mathbf{n} - {}^G d) + n_d$$

Planar Regularities

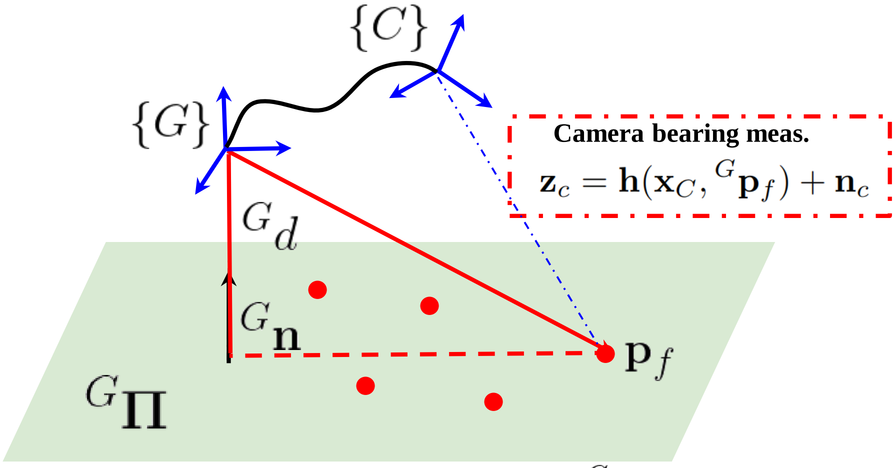


Plane: ${}^G\mathbf{\Pi} = {}^G\mathbf{n}\,{}^G d$ **Point:** ${}^G\mathbf{p}_f$

Point-on-plane regularity

$$z_d = ({}^G\mathbf{p}_f^\top\, {}^G\mathbf{n} - {}^G d) + n_d$$

Planar Regularities



Plane: ${}^G\mathbf{\Pi} = {}^G\mathbf{n}{}^Gd$ **Point:** ${}^G\mathbf{p}_f$

Point-on-plane regularity

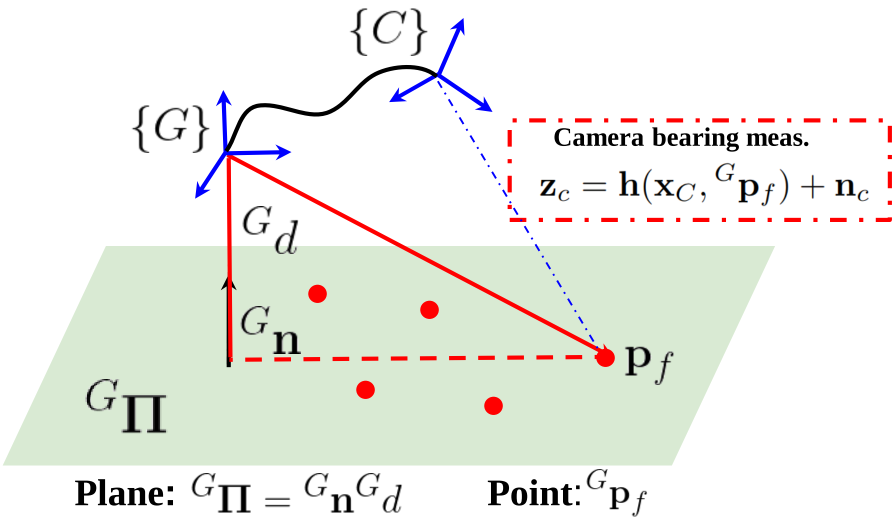
$$z_d = ({}^G\mathbf{p}_f^\top {}^G\mathbf{n} - {}^Gd) + n_d$$

- Stack linearized bearing $\mathbf{z}_C$ and regularity measurements z_d

$$\begin{bmatrix}\tilde{\mathbf{z}}_c \\ \tilde{z}_d\end{bmatrix} = \begin{bmatrix}\mathbf{H}_T^c \\ \mathbf{0}\end{bmatrix}\tilde{\mathbf{x}}_C + \begin{bmatrix}\mathbf{H}_f^c \\ \mathbf{H}_f^d\end{bmatrix}{}^G\tilde{\mathbf{p}}_f + \begin{bmatrix}\mathbf{0} \\ \mathbf{H}_\pi^d\end{bmatrix}{}^G\tilde{\mathbf{\Pi}} + \begin{bmatrix}\mathbf{n}_c \\ n_d\end{bmatrix}$$

$$\Rightarrow \tilde{\mathbf{z}} = \mathbf{H}_T\tilde{\mathbf{x}}_C + \mathbf{H}_f{}^G\tilde{\mathbf{p}}_f + \mathbf{H}_\pi{}^G\tilde{\mathbf{\Pi}} + \mathbf{n}$$

Planar Regularities



Point-on-plane regularity

$$z_d = ({}^G\mathbf{p}_f^\top {}^G\mathbf{n} - {}^G d) + n_d$$

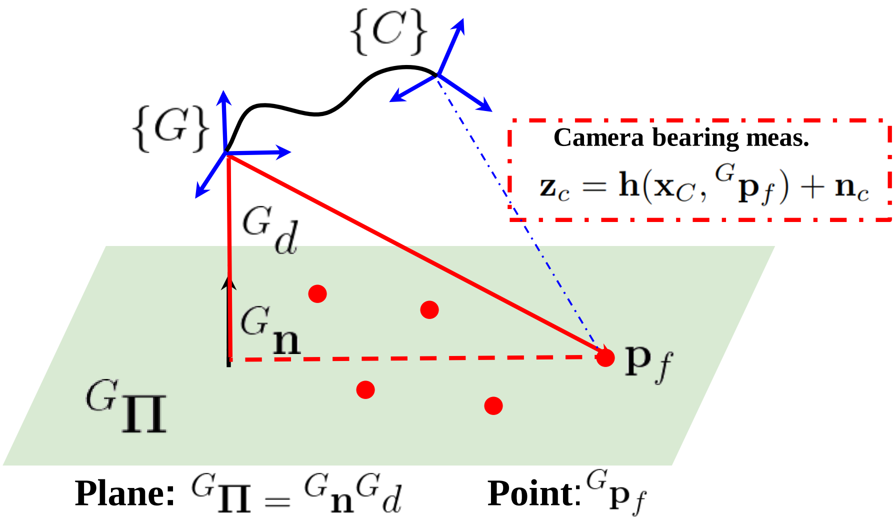
- Stack linearized bearing $\mathbf{z}_C$ and regularity measurements z_d

$$\begin{bmatrix} \tilde{\mathbf{z}}_c \\ \tilde{z}_d \end{bmatrix} = \begin{bmatrix} \mathbf{H}_T^c \\ \mathbf{0} \end{bmatrix} \tilde{\mathbf{x}}_C + \begin{bmatrix} \mathbf{H}_f^c \\ \mathbf{H}_f^d \end{bmatrix} {}^G\tilde{\mathbf{p}}_f + \begin{bmatrix} \mathbf{0} \\ \mathbf{H}_\pi^d \end{bmatrix} {}^G\tilde{\mathbf{\Pi}} + \begin{bmatrix} \mathbf{n}_c \\ n_d \end{bmatrix}$$

$$\Rightarrow \tilde{\mathbf{z}} = \mathbf{H}_T \tilde{\mathbf{x}}_C + \mathbf{H}_f {}^G\tilde{\mathbf{p}}_f + \mathbf{H}_\pi {}^G\tilde{\mathbf{\Pi}} + \mathbf{n}$$

- MSCKF and SLAM feature updates to balance **accuracy** and **efficiency**
- Planar regularities can constrain both in-state and out-of-state features

Planar Regularities



- Stack linearized bearing $\mathbf{z}_c$ and regularity measurements z_d

$$\begin{bmatrix}\tilde{\mathbf{z}}_c \\ \tilde{z}_d\end{bmatrix} = \begin{bmatrix}\mathbf{H}_T^c \\ \mathbf{0}\end{bmatrix}\tilde{\mathbf{x}}_C + \begin{bmatrix}\mathbf{H}_f^c \\ \mathbf{H}_f^d\end{bmatrix}{}^G\tilde{\mathbf{p}}_f + \begin{bmatrix}\mathbf{0} \\ \mathbf{H}_\pi^d\end{bmatrix}{}^G\tilde{\mathbf{\Pi}} + \begin{bmatrix}\mathbf{n}_c \\ n_d\end{bmatrix}$$

$$\Rightarrow \tilde{\mathbf{z}} = \mathbf{H}_T\tilde{\mathbf{x}}_C + \mathbf{H}_f{}^G\tilde{\mathbf{p}}_f + \mathbf{H}_\pi{}^G\tilde{\mathbf{\Pi}} + \mathbf{n}$$

- MSCKF and SLAM feature updates to balance **accuracy** and **efficiency**
- Planar regularities can constrain both in-state and out-of-state features

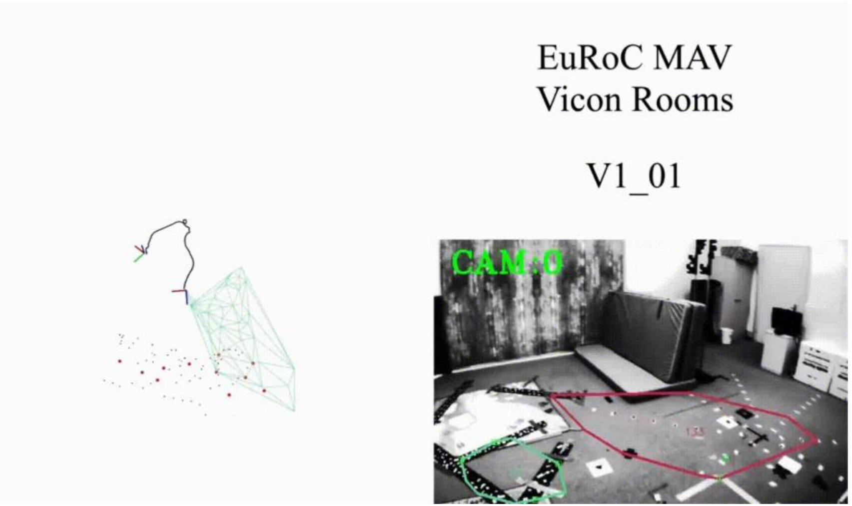
Point-on-plane regularity

$$z_d = ({}^G\mathbf{p}_f^\top {}^G\mathbf{n} - {}^Gd) + n_d$$

SLAM planes shown impressive performance by providing **long-term** tracking and regularization

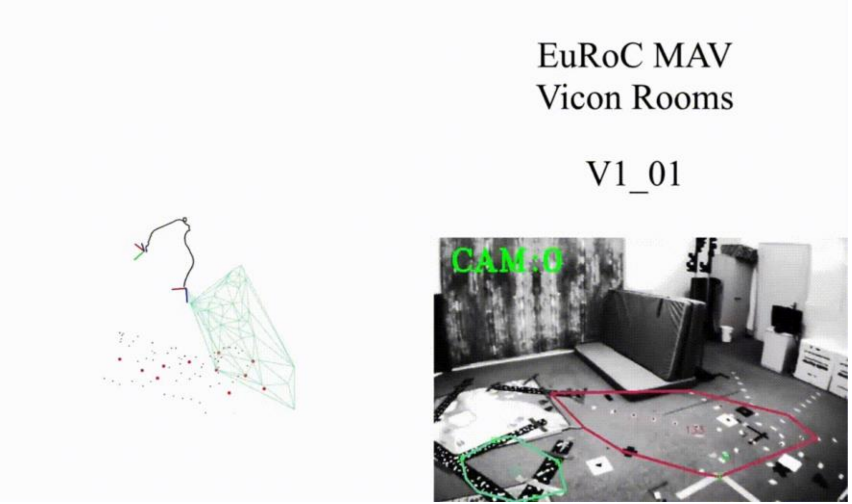
Experimental Results: Detection and tracking

- Planes can be tracked much **longer** than points to better constrain the motion
- Efficient to include planes **without** an additional sensor



Experimental Results: Detection and tracking

- Planes can be tracked much **longer** than points to better constrain the motion
- Efficient to include planes **without** an additional sensor



Dataset	Feat. / PL	PL / Frame	Track Len.	PL Active	Time (ms)
V1_01	19.6 ± 13.3	2.9 ± 1.3	53.4 ± 74.0	0.9 ± 0.7	3.3 ± 0.7
V1_03	10.1 ± 9.4	0.7 ± 1.0	24.9 ± 26.0	0.0 ± 0.2	2.0 ± 0.7
table_01	27.3 ± 13.1	2.7 ± 1.1	61.1 ± 227.6	1.1 ± 0.5	3.5 ± 0.7
table_02	82.0 ± 58.7	2.2 ± 1.3	49.1 ± 249.2	1.2 ± 0.6	4.1 ± 0.9
table_03	33.9 ± 21.3	3.0 ± 1.2	88.5 ± 337.4	1.5 ± 0.6	4.0 ± 0.7

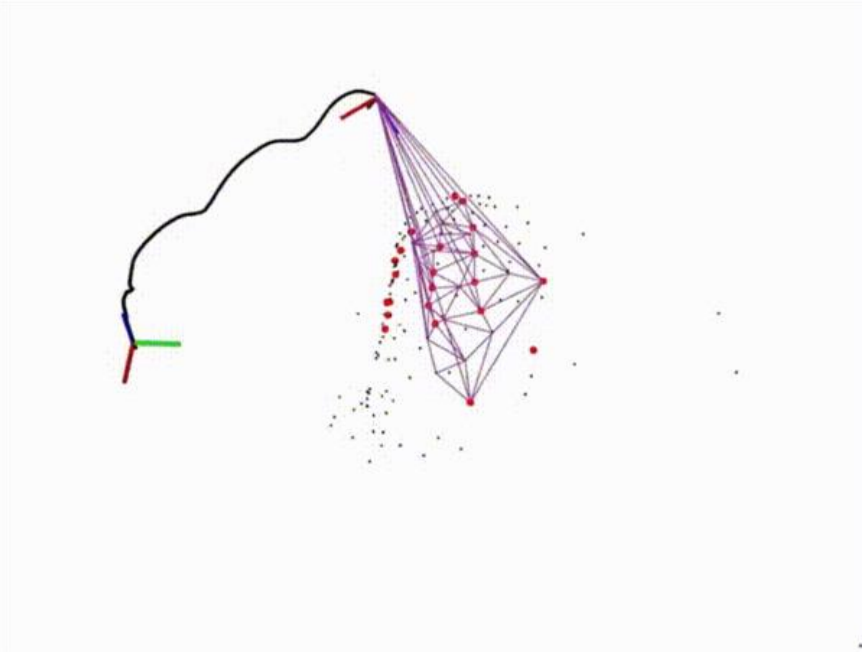
Efficient plane detection and tracking performance!

*All computational results were performed in a single thread on an Intel(R) Xeon(R) E3-1505Mv6 @ 3.00GHz.

Experimental Results

Table 1: Simulation RPE results (degree / cm). **M**: MSCKF feature, **S**: SLAM feature, **PT**: Point, **PL**: Plane.

Algorithm	60m	80m	100m	120m	NEES(3)
M-PT	0.37 / 4.3	0.44 / 5.0	0.50 / 5.6	0.55 / 6.2	3.39 / 1.75
M-PT & M-PL	0.37 / 4.3	0.43 / 4.9	0.48 / 5.5	0.53 / 6.1	3.34 / 1.72
M-PT & MS-PL	**0.36 / 3.6**	**0.42 / 4.1**	**0.48 / 4.6**	**0.53 / 5.1**	3.99 / 1.44
MS-PT	0.30 / 3.6	0.35 / 4.1	0.40 / 4.6	0.43 / 5.1	3.45 / 1.63
MS-PT & M-PL	0.29 / 3.5	**0.33** / 4.0	**0.37** / 4.5	**0.41** / 4.9	3.09 / 1.44
MS-PT & MS-PL	**0.29 / 2.9**	0.35 / **3.3**	0.39 / **3.7**	0.42 / **4.1**	3.38 / 1.20

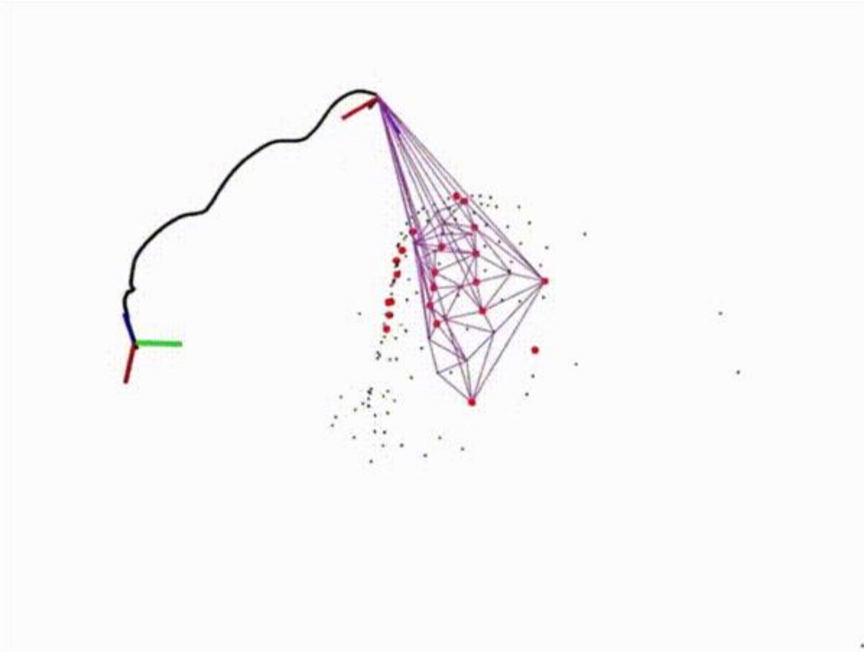


- Impressive performance gain by introducing SLAM planes
- **Consistent** estimation with planar regularities!

Experimental Results

Table 2: Real-world AR Table Dataset (degree / cm)

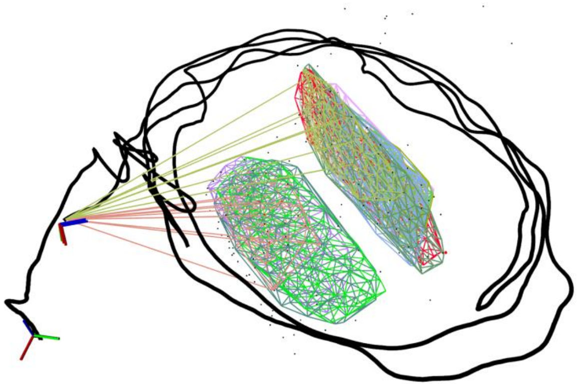
Algorithm	table_01	table_02	table_03	table_04	Time (ms)
M-PT	**0.45** / 6.8	0.85 / 2.4	1.37 / 5.6	0.83 / 7.5	8.7 ± 1.7
M-PT & M-PL	0.52 / 6.5	0.91 / 2.5	1.44 / 5.9	0.87 / 7.1	13.3 ± 3.2
M-PT & MS-PL	0.67 / **4.6**	**0.72** / **2.0**	**0.96** / **3.0**	**0.75** / **3.2**	13.9 ± 2.9
MS-PT	**1.15** / 5.7	1.79 / 4.1	2.41 / 6.9	1.28 / 5.7	9.4 ± 2.0
MS-PT & M-PL	1.32 / 5.5	0.89 / 2.5	**1.03** / **4.5**	1.10 / **4.7**	15.0 ± 3.9
MS-PT & MS-PL	1.25 / **5.1**	**0.65** / **2.3**	1.05 / 4.6	**0.79** / 5.0	14.7 ± 3.2
VINS-Fusion [1]	1.62 / 5.8	1.32 / 3.0	1.47 / 7.6	1.75 / 5.6	35.6 ± 17.0*
OKVIS [2]	2.48 / 9.0	2.01 / 7.7	3.94 / 15.3	2.05 / 16.2	85.5 ± 32.6*



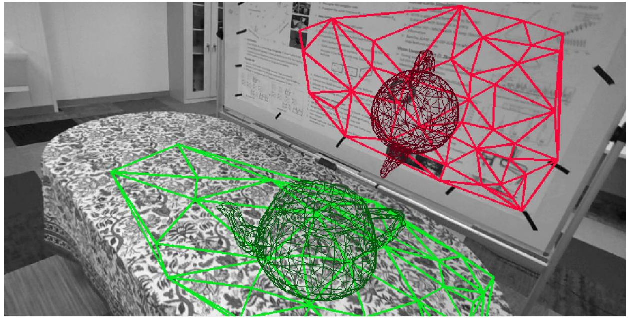
- Including **planes** improves VIO accuracy!
- **Efficient** performance
- Outperform state-of-the-art point-based systems

Monocular Visual-Inertial Odometry with Planar Regularities

Chuchu Chen*, Patrick Geneva*,
Yuxiang Peng, Woosik Lee, and Guoquan Huang



- Novel plane detection and tracking on **monocular** camera
- **Efficient** filter-based VIO with **planar** regularities to improve accuracy
- **Open sourced** code and dataset



RPNG
ROBOT PERCEPTION & NAVIGATION GROUP

ccchu@udel.edu
chuchuchen.net

Source Code

AR Table Dataset