无人系统中的视觉SLAM
-融合环境与载体信息的方法

2019年03月31日浙江大学第二届视觉SLAM研讨会

邹丹平副教授，感知与导航所

上海交通大學
SHANGHAI JIAO TONG UNIVERSITY

面向无人系统的视觉SLAM系统

- 近年来，基于视觉的位姿估算技术发展极快，在众多领域取得应用
- 无人系统 – 无人车、无人机、移动机器人

双目摄像头

背部鱼眼摄像头

双目摄像头

面向无人系统的视觉SLAM 数据集

无人车 – Kitti (2012)

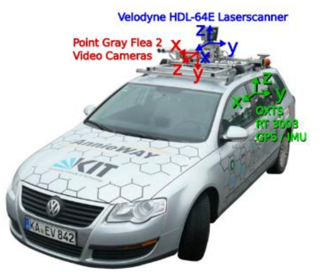
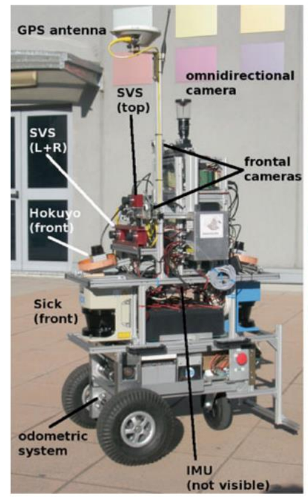
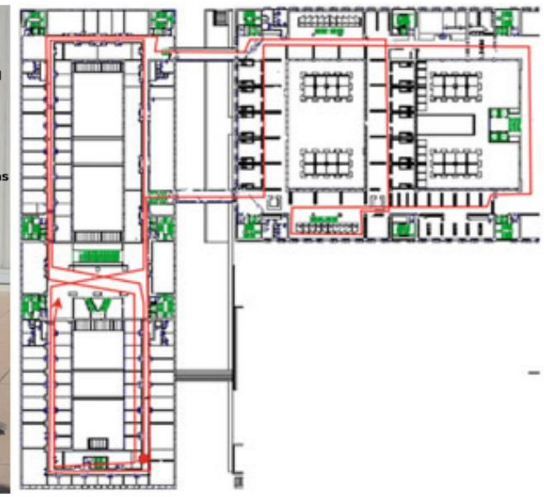


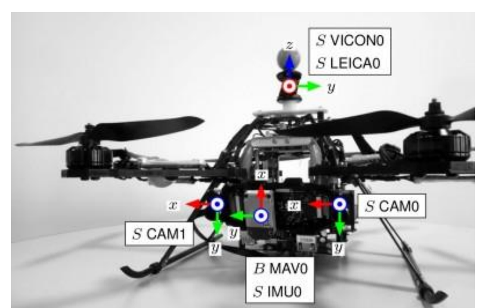
Fig. 1. **Recording Platform.** Our VW Passat station wagon is equipped with four video cameras (two color and two grayscale cameras), a rotating 3D laser scanner and a combined GPS/IMU inertial navigation system.

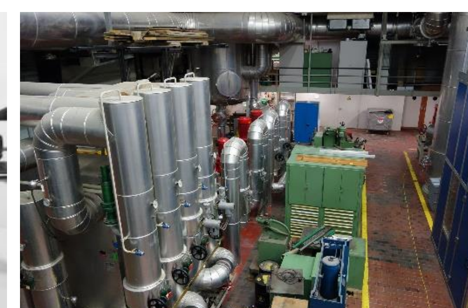
移动机器人 – Rawseeds (2014)



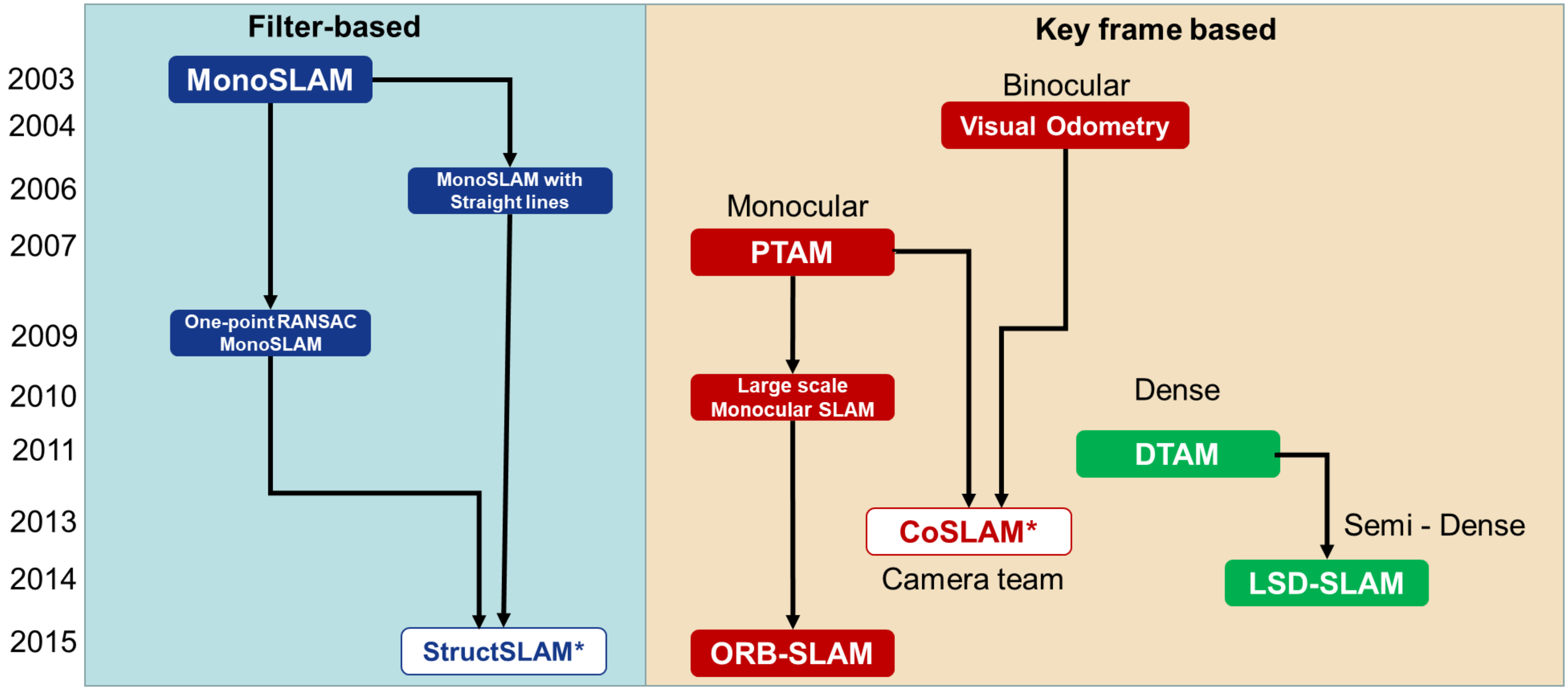


无人机 – Euroc (2016)





视觉SLAM发展历程


Filter-based
Key frame based
2003
2004
2006
2007
2009
2010
2011
2013
2014
2015
MonoSLAM
MonoSLAM with Straight lines
One-point RANSAC MonoSLAM
StructSLAM*
Binocular
Visual Odometry
Monocular
PTAM
Large scale Monocular SLAM
ORB-SLAM
CoSLAM*
Camera team
Dense
DTAM
Semi - Dense
LSD-SLAM

视觉SLAM发展现状

- 理论方法趋向成熟，走向实际应用场景

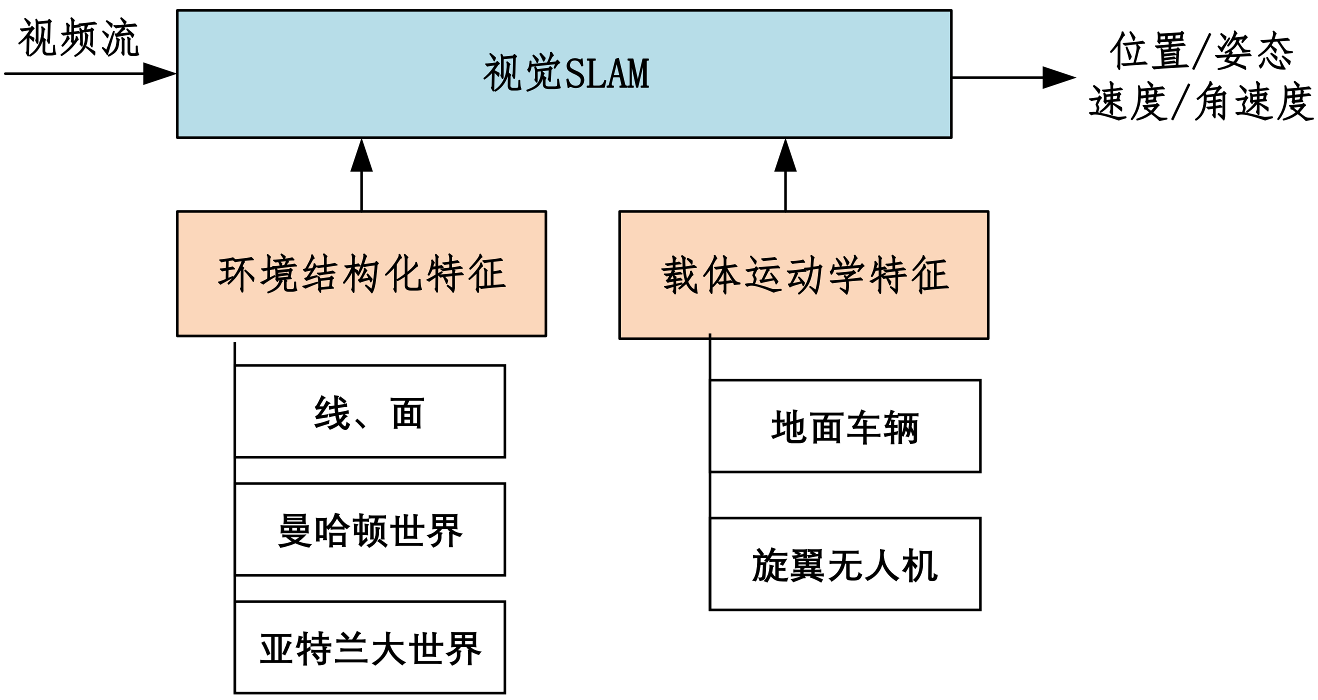
通用视觉SLAM方法 → 面向实际应用场景的高度订制的视觉SLAM方法

视觉SLAM应用场景

- 应用场景要素
  - 环境
    - 高速道路、城市峡谷、商场、居家环境、野外、森林、山洞
  - 载体
    - 二维 - 无人车、服务机器人
    - 三维 – 无人机

融合环境+载体特性的视觉SLAM

- 充分挖掘应用场景中的环境与载体特性，高度订制视觉SLAM系统

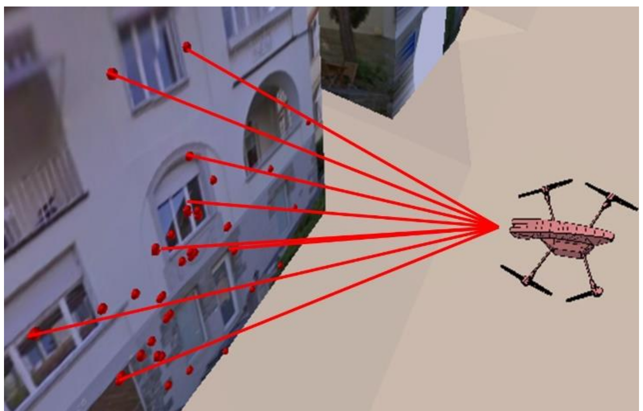


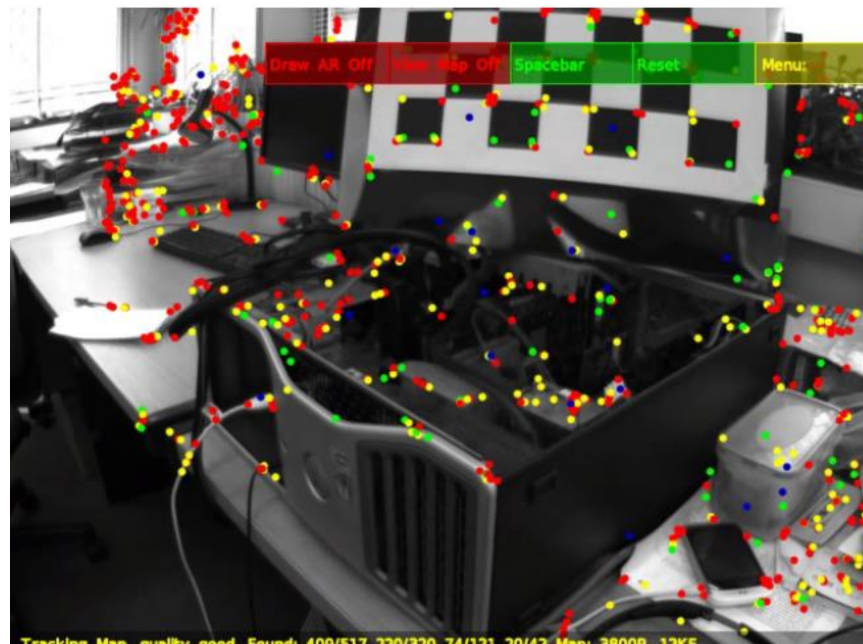
环境特性

- 曼哈顿世界
- 亚特兰大世界－ （多曼哈顿世界）

典型视觉SLAM方法

- 绝大部分方案都基于点特征或者像素特征 - 图像中抽取特征点作为观测值，送入滤波器或者优化框架进行状态估计





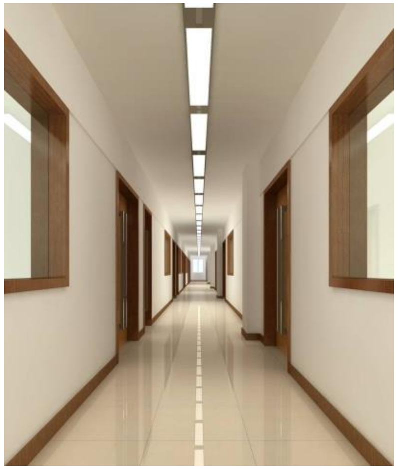
人造环境中的几何结构特性

- 不同于自然环境，人造环境中结构化特性明显

自然场景

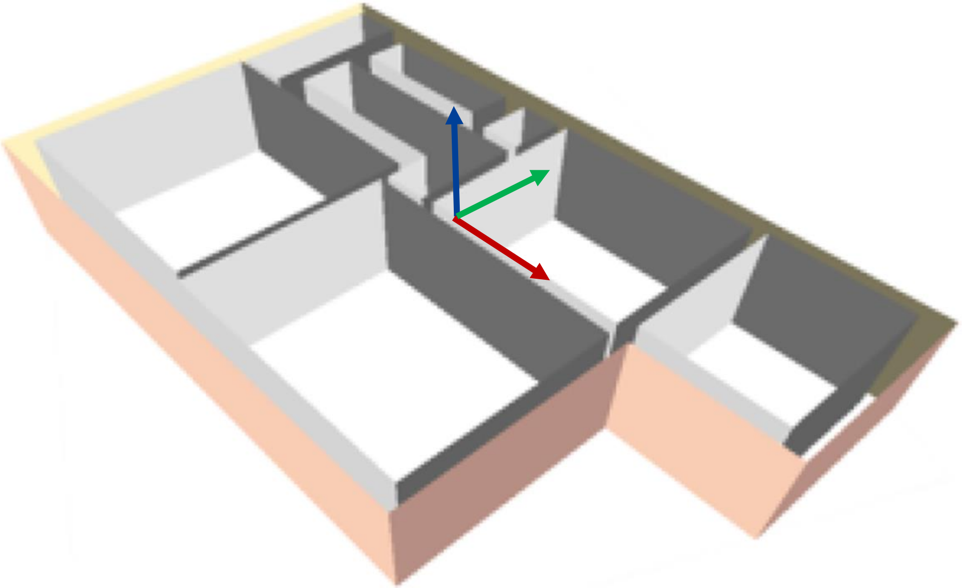
街道场景

地下车库场景



室内场景

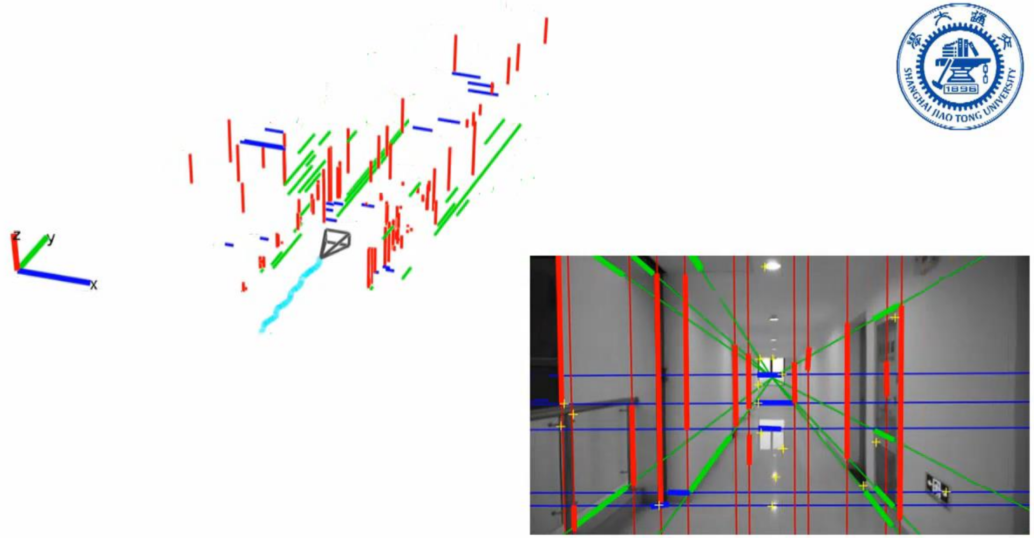
曼哈顿世界假设



1. 线条特征丰富
2. 具有方向性， 三个主方向：x, y, z

基于曼哈顿世界假设的视觉SLAM

- StructSLAM – **结构线条视觉**SLAM
  - 在一些纹理不够丰富的室内场景，线条特征比点特征更加可靠
  - 线条特征的方向性提升了位姿状态的可观测性(observability)



StructSLAM:Visual SLAM with Building Structure Lines

Zhou, Huizhong, Danping, Zou, et al. "StructSLAM: Visual SLAM with building structure lines." *Vehicular Technology, IEEE Transactions on* 64.4 (2015): 1364-1375. - Special session for indoor localization

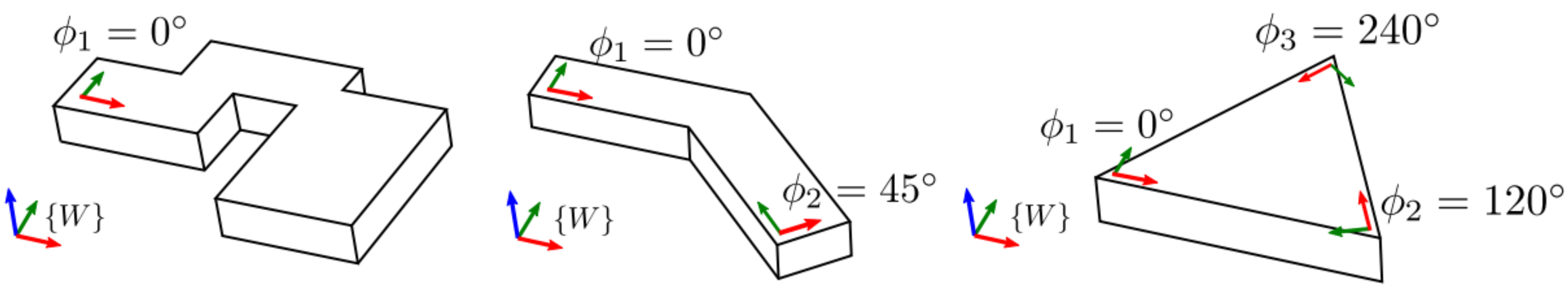
曼哈顿世界假设

- 建筑存在多样性

亚特兰大世界-多曼哈顿世界假设

- 将一个复杂的人造环境用多个局部曼哈顿世界逼近

  每个局部曼哈顿世界由一个水平方向(ϕ)确定



一个曼哈顿世界　　　两个曼哈顿世界　　　三个曼哈顿世界

基于亚特兰大世界假设的视觉SLAM

- **基于亚特兰大世界假设的视觉惯性SLAM：** StructVIO
- 主要贡献：
  - 1. 提出使用**多个曼哈顿世界**叠加的模型，处理更加不规则的场景
  - 2. 采用**多状态滤波的视觉+惯性**紧耦合数学框架，并对特征处理进行进一步改进
  - 3. 一个室内外视觉惯性SLAM公开数据集

Zou, Danping, et al. "StructVIO: Visual-inertial Odometry with Structural Regularity of Man-made Environments." *arXiv preprint arXiv:1810.06796* (2018). (Accepted by IEEE Trans. on Robotics)

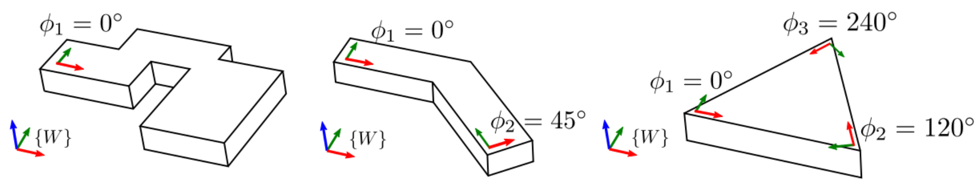
项目网址：http://drone.sjtu.edu.cn/dpzou/project/structvio.html

StructVIO - 技术细节

- 亚特兰大世界中结构线条表示
  - **线条参数表示法**
  - **结构线条的投影方程**
- 多状态紧耦合滤波框架设计
  - **状态定义**
  - **多状态观测方程**
- 滤波器前端
  - **结构线条初始化与三角化更新**
  - **特征先验累积**
  - 局部曼哈顿世界管理
  - 结构线条检测与跟踪
  - 结构线条分类
  - Outlier去除

坐标系定义

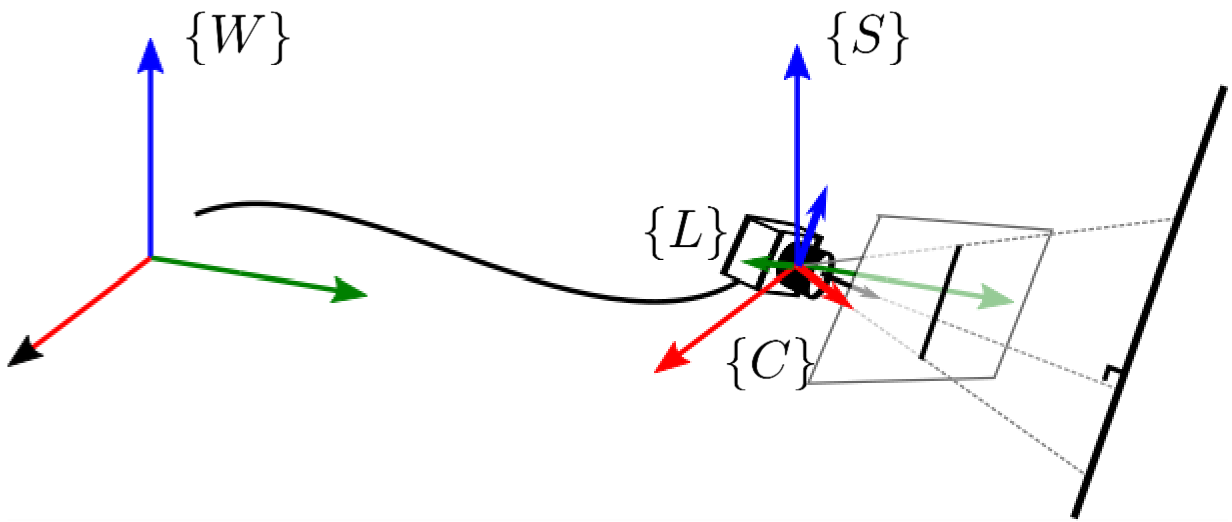
- 世界坐标系 $\{W\}$
  - Z轴与重力反方向一致
  - 坐标原点在起始点



- 局部曼哈顿世界 $\phi_i \in [0, \pi/2)\,, i = 1, \ldots, N$
- 相机坐标系 $\{C\}$
  - Z轴方向与相机观察方向一致
  - X、Y轴方向与图像x，y轴保持一致
- 起始坐标系 $\{S\}$ - 滑动曼哈顿世界坐标系
  - 原点与相机坐标系一致
  - 坐标轴与局部曼哈顿世界一致

亚特兰大世界中结构线条表示

- 采用以相机为中心(Camera-centric)的特征表达方式



- 结构线条参数空间 - $\{L\}$ - 以相机光学中心为原点的坐标系,Z轴方向与线条方向保持一致

亚特兰大世界中结构线条表示

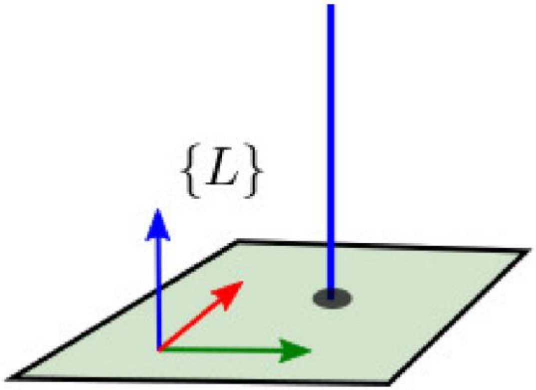
- 参数空间$\{L\}$中，一根结构线条可表示XY平面上的一个点表示：

$$^{L}l_p = (a, b, 0)^{\mathrm{T}}$$

- 为了达到更好的线性化效果，我们采用反深度表示 $^{L}l_p$

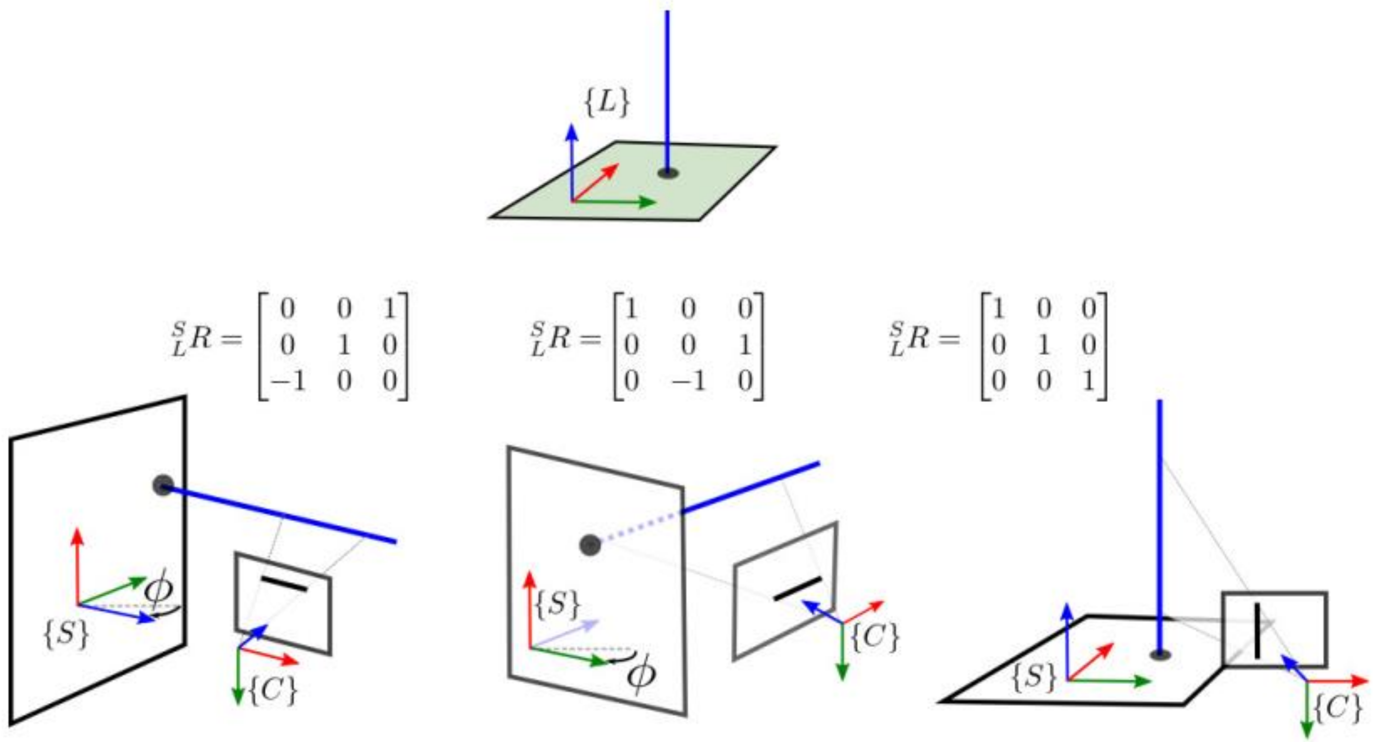
$$^{L}l_p = (\theta, \rho, 0)^{\mathrm{T}}$$

$$\theta = \mathrm{atan2}(b, a)$$
$$\rho = 1/\sqrt{a^2 + b^2}$$



亚特兰大世界中结构线条表示

- 通过旋转矩阵${}_{L}^{S}R$ 将参数空间的直线映射到局部曼哈顿世界的三个主方向：



亚特兰大世界下结构线条表示

- 通过水平旋转，将结构线条旋转从局部曼哈顿世界(ϕ_i)映射到全局世界坐标系

$$
{}^{W}_{S}R(\phi_i) = \begin{bmatrix} \cos(\phi_i) & \sin(\phi_i) & 0 \\ -\sin(\phi_i) & \cos(\phi_i) & 0 \\ 0 & 0 & 1 \end{bmatrix}
$$

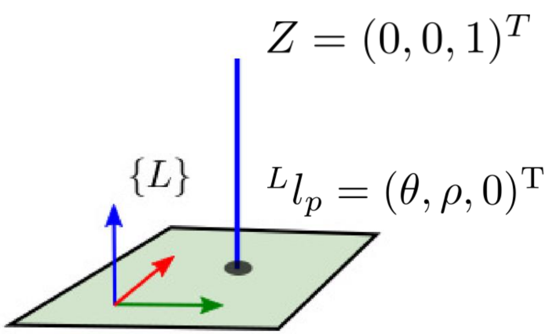
- 再通过世界坐标系$\{W\}$到相机坐标系$\{C\}$的变换，

$$
{}^{C}_{W}\tau = ({}^{C}_{W}R, {}^{C}p_W)
$$

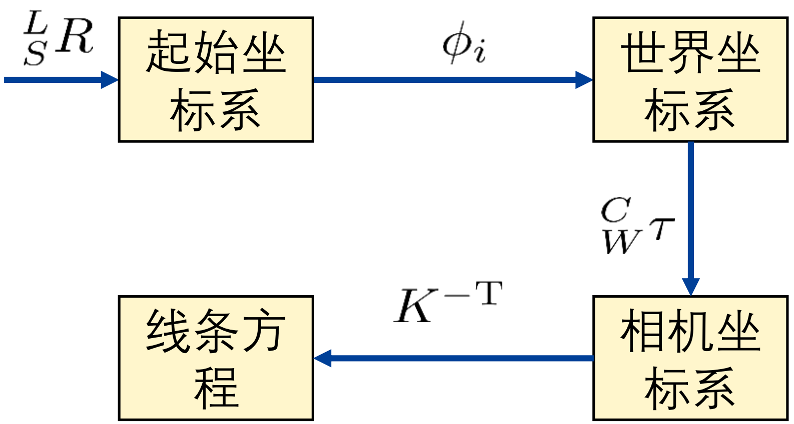
转到当前相机坐标系。

结构线条的投影

- 将参数空间中的Z轴方向与交点分别进行坐标转换



线条参数空间



$$ {}^{im}l = (K^{-T})({}^{C}l_p \times {}^{C}v) \qquad {}^{C}l_p,\ {}^{C}v $$

结构线条的投影

- 结构线条投影方程:

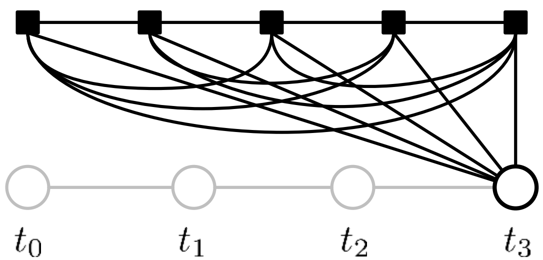
$$^{im}l = \Pi(l, \phi_i, {}^S_L R, {}^W_C \tau)$$

$$^{im}l = \Pi(l, \phi_i, {}^S_L R, {}^I_C \tau, {}^W_I \tau)$$

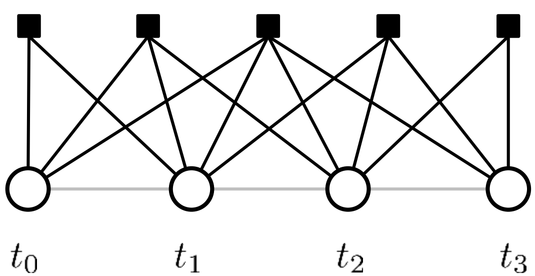
- 其中${}^S_L R$ 为常量（在识别线条方向之后）
- 根据投影方程设计滤波器，同时更新$l, \phi_i, {}^W_I \tau$

视觉+惯性的紧耦合方法

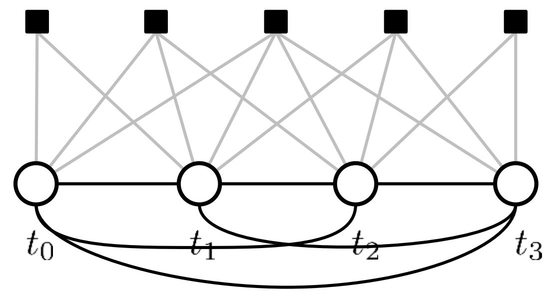
- EKF-SLAM方法、多状态约束方法和关键帧优化方法对比示意图，其中黑色方框表示特征点，圆圈表示在某时刻的相机状态



经典滤波方法



关键帧优化方法



多状态约束滤波方法

多状态约束紧耦合方法

- **相比经典滤波方法：**
  - 特征点三维坐标解耦，将特征点状态估算排除在滤波器之外，降低了状态维度，提升了处理大量特征点的能力。
- **相比关键帧优化方法：**
  - 逐帧更新和边缘化，充分利用了持续时间较短的特征点轨迹（机会轨迹）
  - 可单条轨迹更新，增强了灵活性，方便控制计算峰值
  - 达到相近精度情况下，多状态约束方法计算成本较低。

多状态约束紧耦合滤波

- 采用多状态约束滤波框架, 状态量定义如下

$$x_k = [x_{I_k}, {}^{I}_{C}\tau, \boxed{\phi_1, \cdots \phi_N}, \boxed{{}^{W}_{I_1}\tau, \cdots {}^{W}_{I_M}\tau}]$$

$$x_{I_k} = [{}^{W}_{I_k}\tau, {}^{W}v_{I_k}, b_{g_k}, b_{a_k}]$$

当前时刻IMU状态

IMU与相机之间的相对位姿

局部曼哈顿世界方向

历史IMU位姿

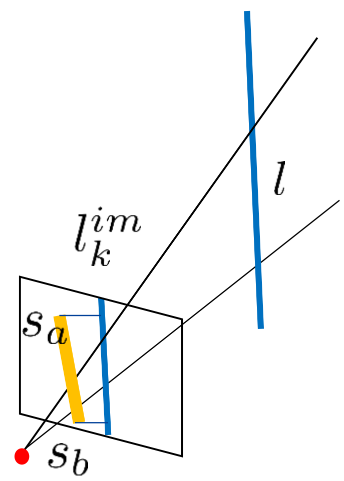
多状态约束紧耦合滤波

- **解耦** - 将观测模型从重投影误差变为位姿约束，消除滤波器对结构线条参数的依赖
- **重投影误差**观测模型：
  - k 时刻下结构线条 l 在图像中的投影：

$$l_k^{im} = \Pi(l, \phi, {}_L^S R, {}_C^W \tau)$$

  - 图像上所观测的对应线段为 $s_a \leftrightarrow s_b$
  - 重投影误差为两个短点到线条投影的直线距离（带符号）

$$\begin{aligned} r_k &= D(s_a, s_b, l_k^{im}) \\ &= D(s_a, s_b, \Pi(l, \phi, {}_L^S R, {}_C^W \tau)) \end{aligned}$$



多状态约束紧耦合滤波

- EKF滤波的局部线性近似的观测模型

$$r_k = h_0 + J_l \delta l + J_\phi \delta\phi + J_{IC} \delta\, {}^I_C\tau + J_{WI_k} \delta\, {}^W_{I_k}\tau$$

- 同一根结构线条从1到M时刻所有的观测模型

$$\begin{aligned} r_1 &= h_0 + J_l \delta l + J_\phi \delta\phi + J_{IC} \delta\, {}^I_C\tau + J_{WI_1} \delta\, {}^W_{I_1}\tau \\ &\cdots \quad \cdots \\ r_k &= h_0 + J_l \delta l + J_\phi \delta\phi + J_{IC} \delta\, {}^I_C\tau + J_{WI_k} \delta\, {}^W_{I_k}\tau \\ &\cdots \quad \cdots \\ r_M &= h_0 + J_l \delta l + J_\phi \delta\phi + J_{IC} \delta\, {}^I_C\tau + J_{WI_M} \delta\, {}^W_{I_M}\tau \end{aligned}$$

$$\Downarrow$$

$$z = H_l \underbrace{\boxed{\delta l}}_{\text{线条参数}} + H_\phi \underbrace{\boxed{\delta\phi}}_{\text{局部曼哈顿世界方向}} + H_{CI} \underbrace{\boxed{\delta\, {}^I_C\tau}}_{\text{相机IMU相对变换}} + H_{WI} \underbrace{\boxed{[\delta\, {}^W_{I_1}\tau \cdots \delta\, {}^W_{I_M}\tau]}}_{\text{1~M时刻的IMU位姿}}$$

多状态约束紧耦合滤波

- 通过将残差z往H_l左零空间投影，消除对结构线条参数δl的依赖：

$$z = H_l \delta l + H_\phi \delta \phi + H_{CI} \delta\, {}^I_C\tau + H_{WI}[\delta^W_{I_1}\tau \cdots \delta^W_{I_M}\tau]$$

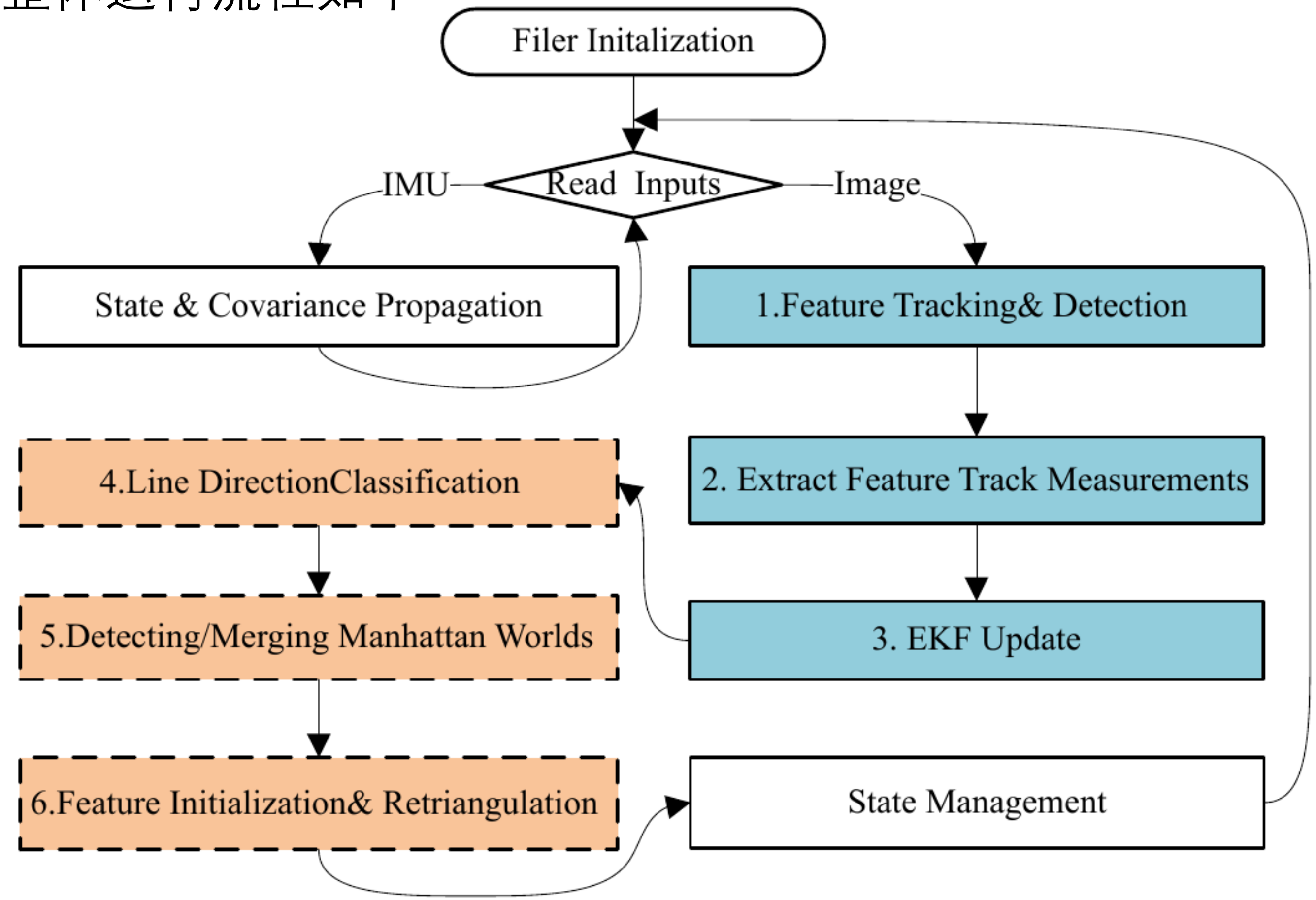
$$\Downarrow$$

$$z^{(0)} = H^{(0)}_\phi \delta\phi + H^{(0)}_{CI} \delta\, {}^C x_I + H^{(0)}_{WI}[\delta^W_{I_1}\tau \cdots \delta^W_{I_M}\tau],$$

- 观测方程与
  - 1. 局部曼哈顿世界朝向
  - 2. IMU与摄像头的相对位姿
  - 3. IMU的历史位姿（该结构线条可见时间范围）

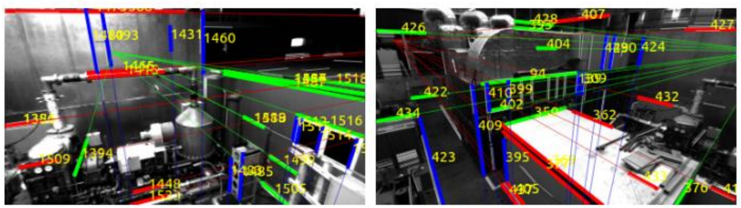
多状态约束紧耦合滤波

- 算法整体运行流程如下

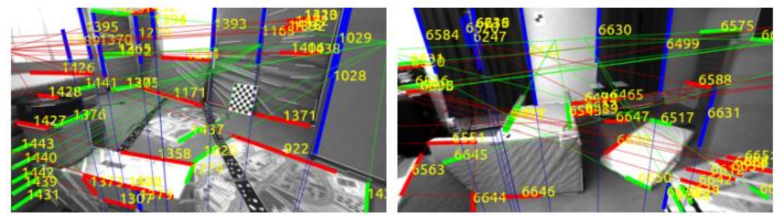


实验结果

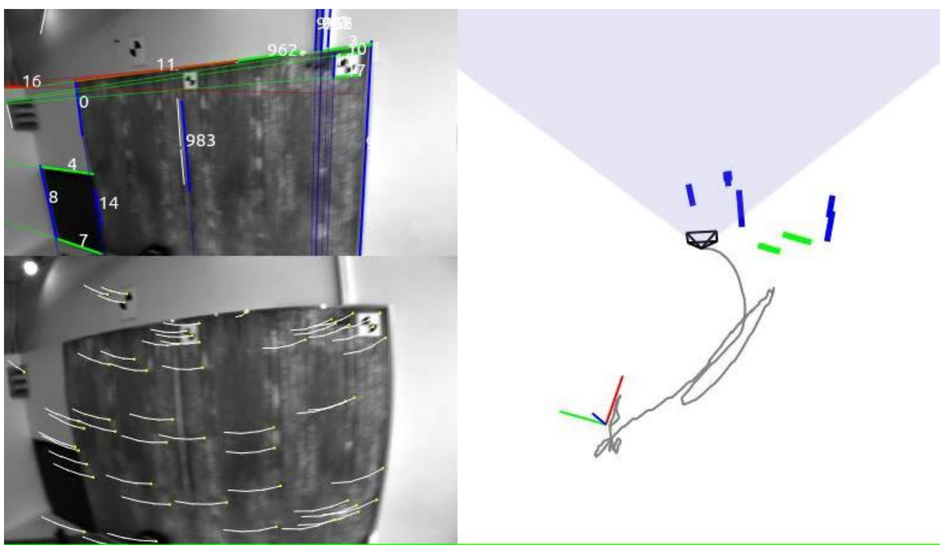
- 公开数据集测试（Euroc）



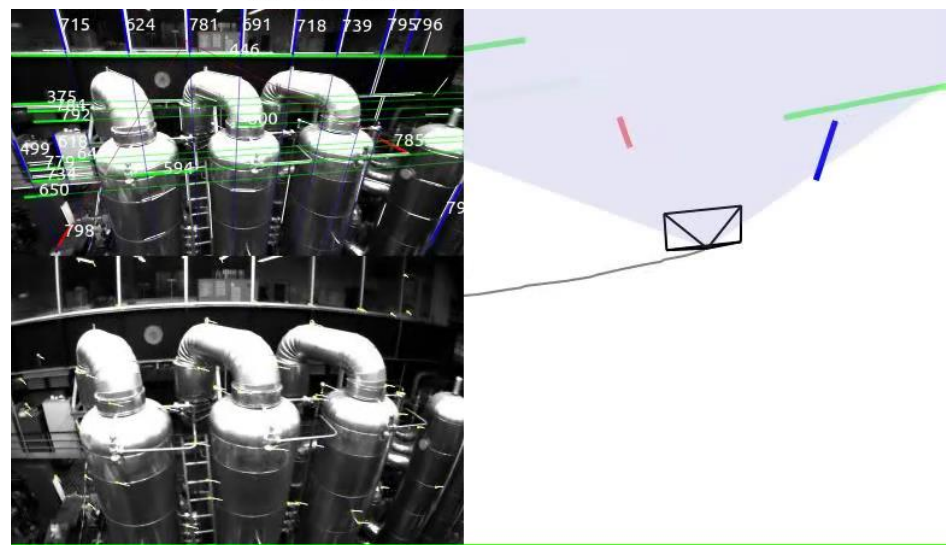
(a) Machine hall



(b) Vicon room



V2_03_difficult



MH_05_difficult

实验结果

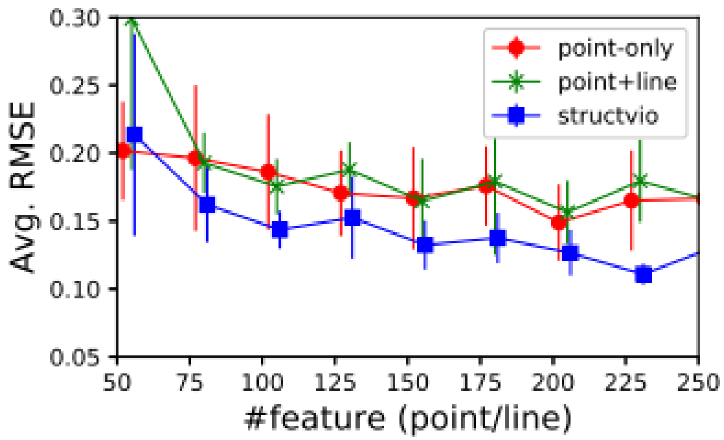
- 公开数据集测试（Euroc）

Dataset	OKVIS[5]		VINS[6](w/o loop)		StructVIO	
	RMSE	Max.	RMSE	Max.	RMSE	Max.
MH_01_easy	0.308	0.597	**0.157**[2]	0.349	**0.079**[1]	0.251
MH_02_easy	0.407	0.811	**0.181**[2]	0.533	**0.145**[1]	0.267
MH_03_medium	0.241	0.411	**0.196**[2]	0.450	**0.103**[1]	0.271
MH_04_difficult	0.363	0.641	**0.345**[2]	0.475	**0.130**[1]	0.286
MH_05_difficult	0.439	0.751	**0.303**[2]	0.434	**0.182**[1]	0.358
V1_01_easy	**0.076**[2]	0.224	0.090	0.201	**0.060**[1]	0.180
V1_02_medium	0.141	0.254	**0.098**[1]	0.334	**0.130**[2]	0.260
V1_03_difficult	0.240	0.492	**0.183**[2]	0.376	**0.090**[1]	0.263
V2_01_easy	0.134	0.308	**0.080**[2]	0.232	**0.045**[1]	0.140
V2_02_medium	0.187	0.407	**0.149**[2]	0.379	**0.066**[1]	0.157
V2_03_difficult	**0.255**[2]	0.606	0.268	0.627	**0.110**[1]	0.231

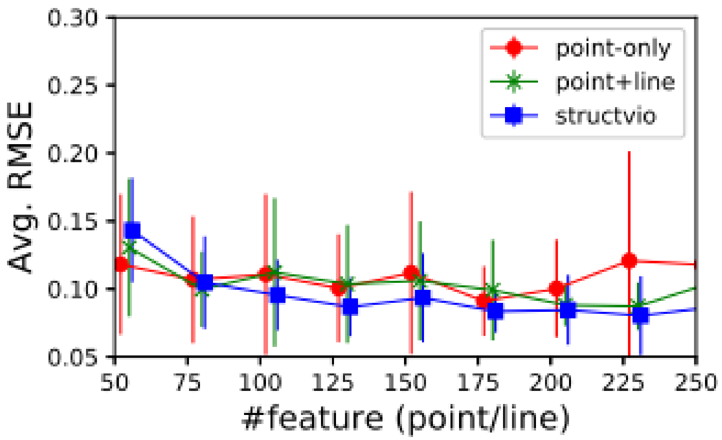
RMSE-Rooted Mean Squared Error

实验结果

- 公开数据集测试（Euroc）
  - Machine hall有更强的结构性，StructVIO性能提升明显



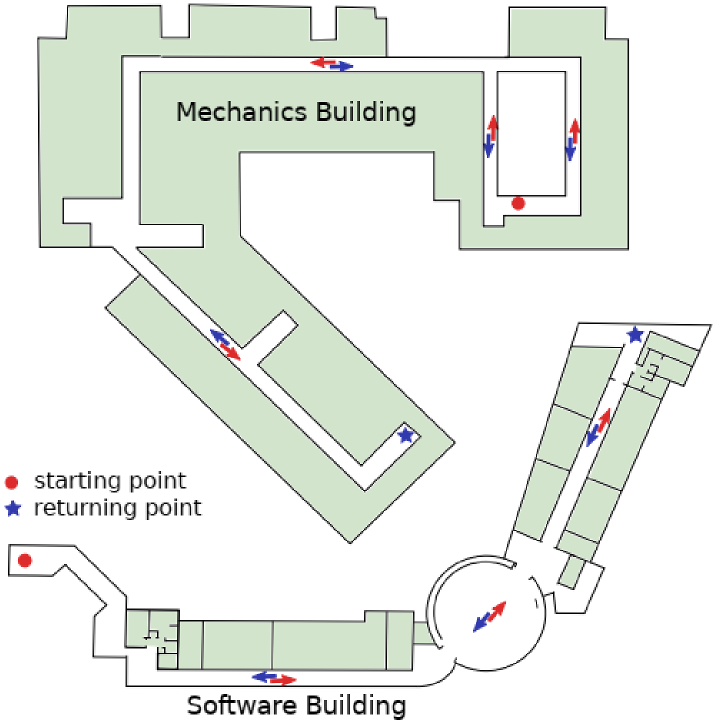
(a) Machine hall



(b) Vicon room

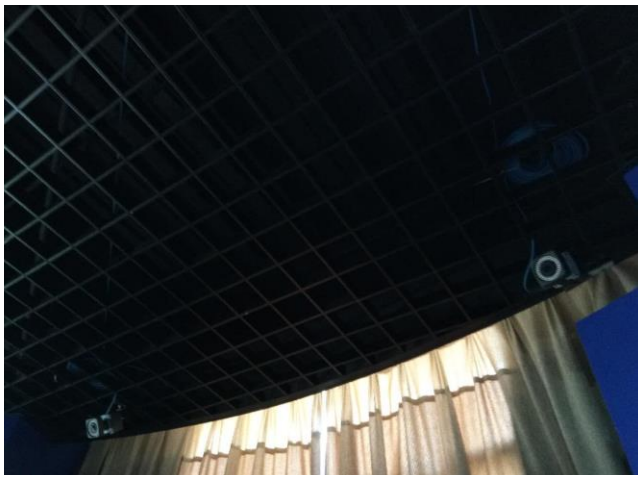
实验结果

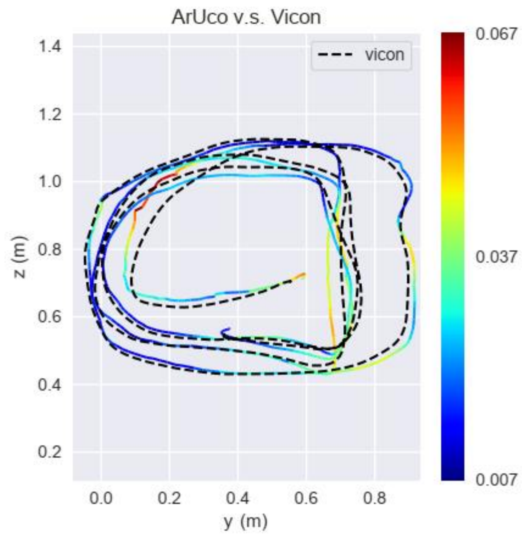
- 使用Google Tango平板自采数据集（16组），微电子楼、软件学院、机动学院、每组行走5~10分钟，覆盖室内外，光照变化剧烈。



实验结果

- 真值部分采用vicon动作捕获系统，部分采用二维码





起始时间段：s

结束时间段：e

对齐起始段：$T^{*}=\arg\min_{T}\sum_{t\in s}(\|T(p^{t})-g_{s}^{t}\|^{2})$

计算结束段RMSE与Max误差：

$$RMSE=\sqrt{\frac{1}{|e|}\sum_{t\in e}\|(T(p^{t})-g_{e}^{t}\|^{2})}\qquad Max.=\max|T(p^{t})-g_{e}^{t}|$$

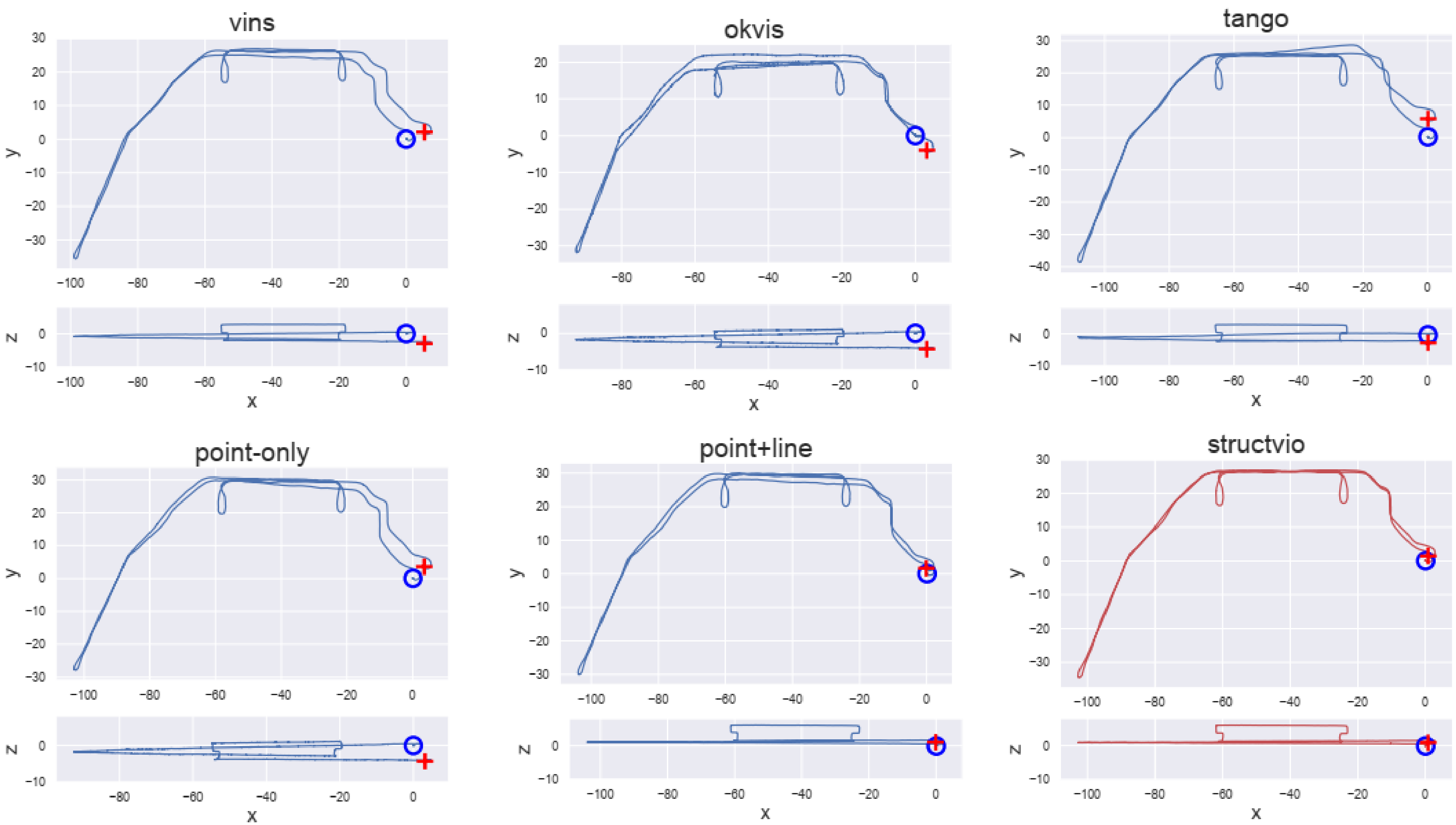
实验结果

- 比较方法：OKVIS, VINS, Project Tango, Point-only, Point-line, StructVIO

Seq. Name	Traveling Dist. [m]	OKVIS[5]		VINS[6](w/o loop)		Project Tango		Point-only		Point+Line		StructVIO	
		RMSE	Max.	RMSE	Max.	RMSE	Max.	RMSE	Max.	RMSE	Max.	RMSE	Max
Soft-01	315.167	6.702	9.619	4.861	6.688	5.715	8.181	**2.153**[2]	2.728	2.262	2.842	**1.931**[1]	2.437
Soft-02	438.198	4.623	6.713	2.713	4.086	4.238	6.226	3.905	5.243	**1.468**[2]	2.026	**1.429**[1]	1.984
Soft-03	347.966	**4.505**[2]	6.223	7.270	9.832	167.825	228.630	6.515	8.119	8.618	10.790	**0.325**[1]	1.020
Soft-04	400.356	3.993	5.784	28.667	75.479	2.453	3.544	**1.550**[1]	2.028	4.051	5.262	**1.722**[2]	2.241
Mech-01	340.578	3.627	4.745	2.452	3.260	**1.948**[2]	2.726	3.298	3.961	4.323	5.181	**0.909**[1]	1.165
Mech-02	388.548	3.079	4.195	3.570	4.754	**1.596**[2]	2.217	1.663	2.108	2.317	2.927	**0.779**[1]	1.022
Mech-03	317.974	3.875	5.324	4.682	9.113	4.220	5.781	**2.384**[2]	3.020	4.193	5.272	**1.161**[1]	1.532
Mech-04	650.430	-	-	3.002	8.592	1.915	5.808	1.785	4.663	**1.425**[2]	3.729	**0.742**[1]	1.940
MicroA-01	257.586	2.485	3.382	**0.654**[2]	1.148	45.599	61.058	2.849	3.505	2.189	2.721	**0.642**[1]	1.225
MicroA-02	190.203	3.428	5.186	14.222	57.172	**1.145**[1]	1.692	1.964	2.514	**1.723**[2]	2.207	2.089	2.661
MicroA-03	388.730	0.078	0.779	**1.800**[1]	2.578	4.400	6.253	3.824	5.169	3.072	4.232	**1.884**[2]	2.892
MicroA-04	237.856	6.136	8.532	**0.994**[2]	1.765	55.200	75.318	2.056	2.897	2.406	2.879	**0.350**[1]	0.448
MicroB-01	338.962	2.898	4.025	**1.856**[2]	2.944	38.197	50.572	7.084	8.576	7.337	8.913	**1.477**[1]	1.902
MicroB-02	306.316	2.240	3.490	**1.030**[2]	2.431	5.660	8.652	2.521	3.714	3.197	4.610	**0.470**[1]	0.799
MicroB-03	485.291	-	-	2.132	3.368	**2.009**[2]	2.960	6.490	8.978	4.507	6.301	**0.445**[1]	0.675
MicroB-04	357.251	4.064	6.481	**1.332**[2]	2.068	13.962	22.028	5.078	7.713	1.977	3.074	**0.473**[1]	0.777
Mean Drift Err.(%)		1.078%		1.410%		6.180%		0.957%		**0.956%**[2]		**0.292%**[1]	
Median Drift Err.(%)		0.781%		**0.538%**[2]		0.900%		0.559%		0.570%		**0.176%**[1]	

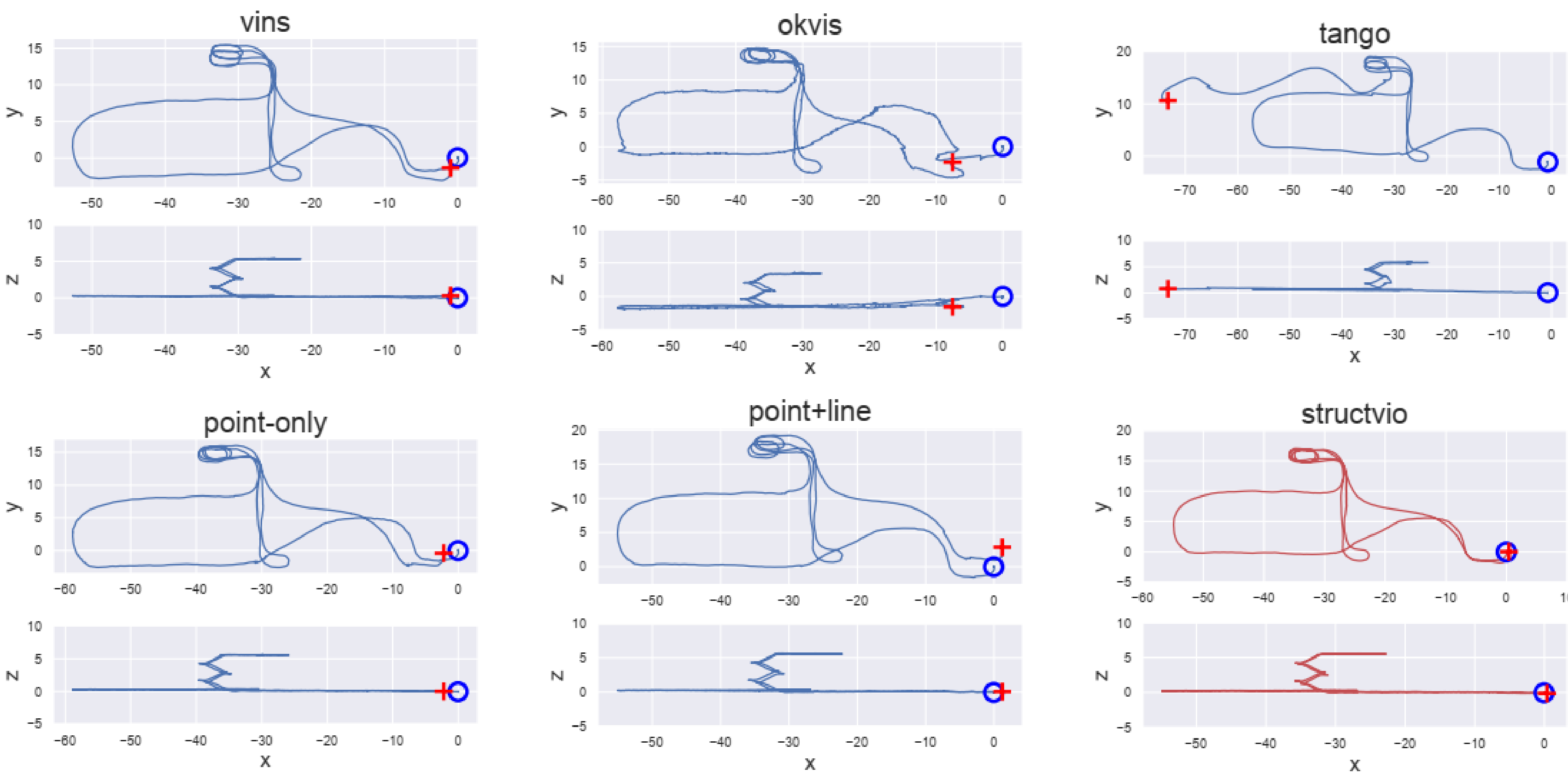
实验结果

- 软件学院场景(Soft-02)



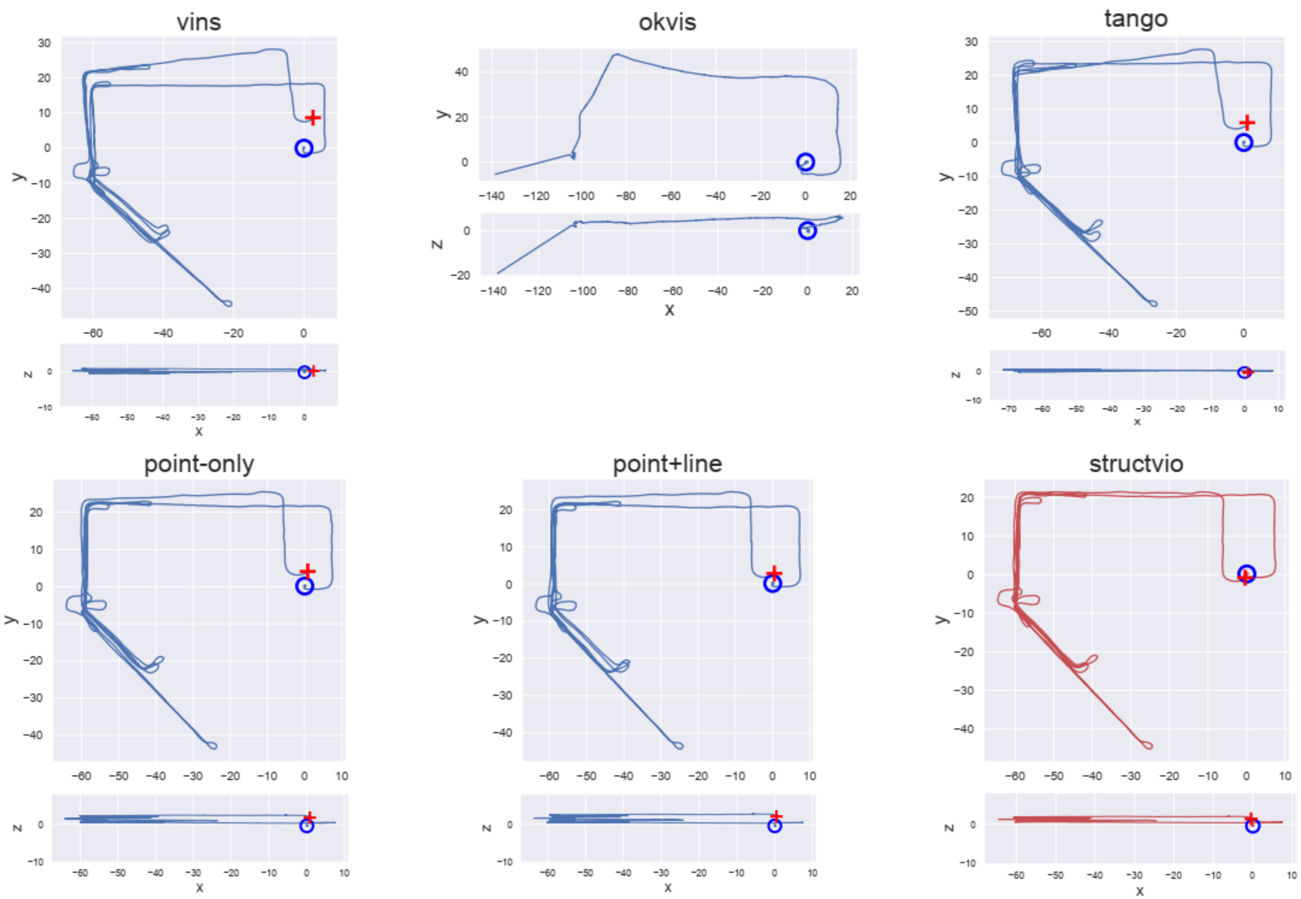
实验结果

- 微电子楼场景(MicroA-04)



实验结果

- 机动学院场景 (Mech-04)

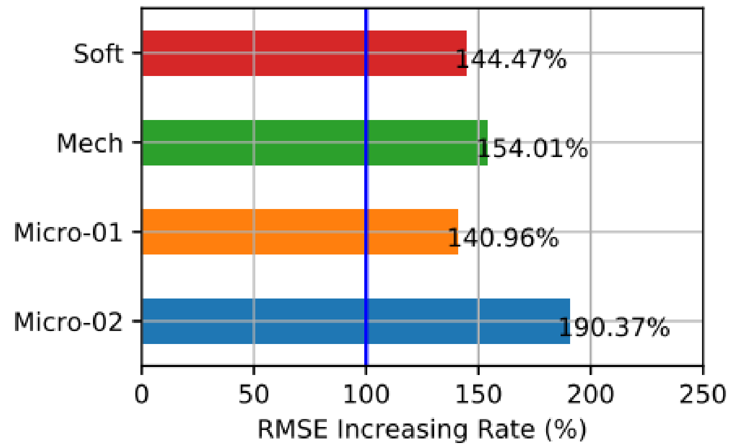


实验结果

- 其它测试

Seq. Name	Atlata world		Manhattan world	
	RMSE	Max.	RMSE	Max.
Mech-01	0.909	1.165	1.144	1.524
Mech-02	0.779	1.022	1.286	1.061
Mech-03	1.161	1.532	2.029	1.211
Mech-04	0.742	1.940	1.822	2.193
Soft-01	1.931	2.437	2.896	2.397
Soft-02	1.429	1.984	3.092	4.149
Soft-03	0.325	1.020	3.352	4.236
Soft-04	1.722	2.241	3.178	4.120

曼哈顿世界vs多曼哈顿世界



未使用特征信息累积

- 论文地址 (https://arxiv.org/abs/1810.06796)
- 软件以及数据集已在网上公布

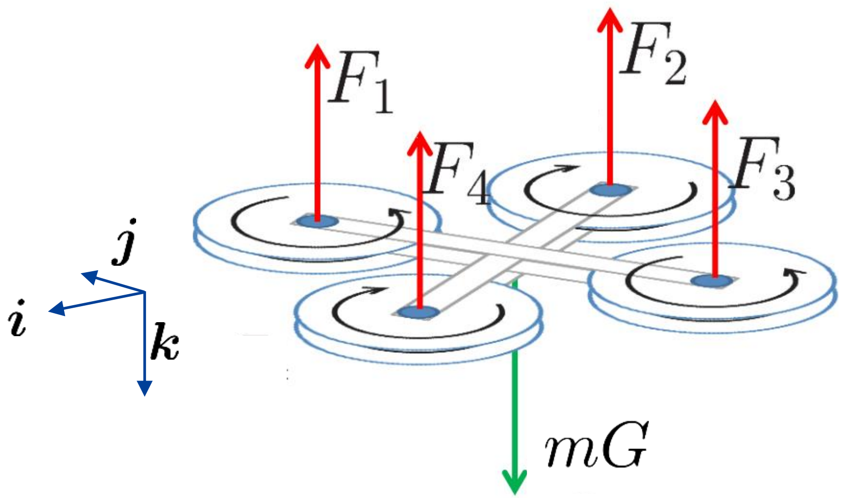
http://drone.sjtu.edu.cn/dpzou/project/structvio/structvio.html

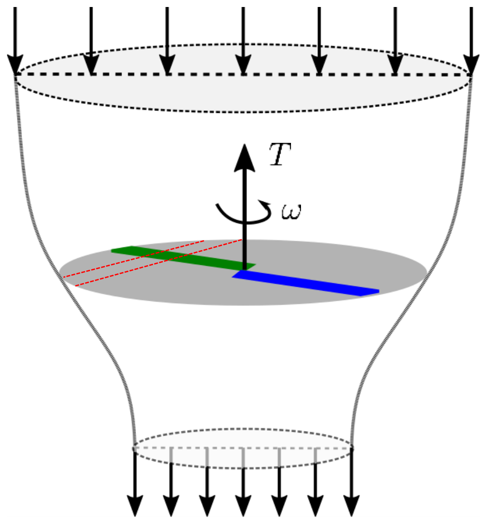
载体特性

- 车辆运动特性
  - 轮速计
    - 速度与水平朝向信息
  - 非完整性运动约束
    - 体坐标系下侧向与垂直方向速度为零
- 旋翼无人机空气动力学特性

旋翼无人机运动特性

- 旋翼无人机动力学原理

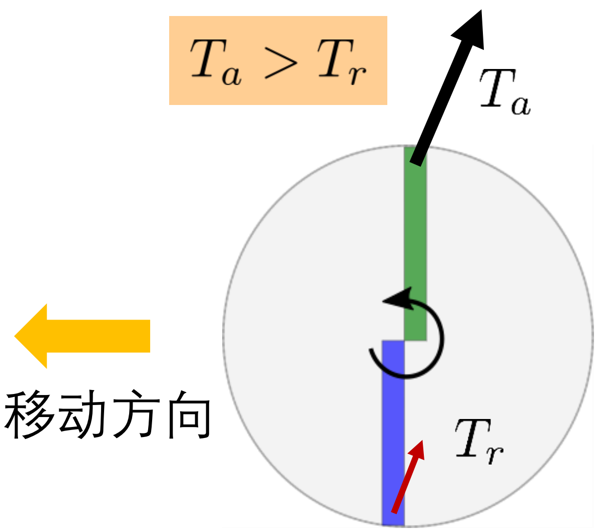


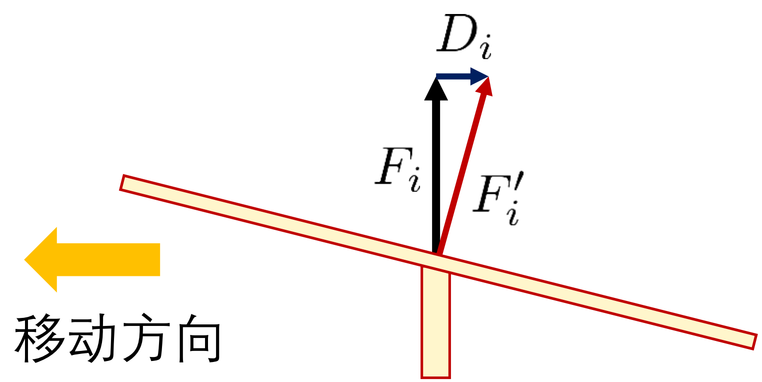


$$F_i = \alpha\omega_i^2$$

旋翼无人机运动特性

- Blade flapping现象 – 螺旋桨平移时，在螺旋桨平面收到额外的阻力





Wang, Rongzhi, Danping Zou, etc. "An aerodynamic model-aided state estimator for multi-rotor uavs." In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 2164-2170. IEEE, 2017.

Sartori, Daniele, Danping Zou,etc. "A Revisited Approach to Lateral Acceleration Modeling for Quadrotor UAVs State Estimation." In 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 5711-5718. IEEE, 2018.

旋翼无人机运动特性

- 在体坐标系下(Body frame)，每个螺旋桨在螺旋桨平面所受的额外阻力与移动方向相反，可由如下式子进行近似计算

$$D_i \approx -\lambda\omega_i \boldsymbol{v}^{\perp}$$

常量

（重心）移动速度在螺旋桨平面中的投影

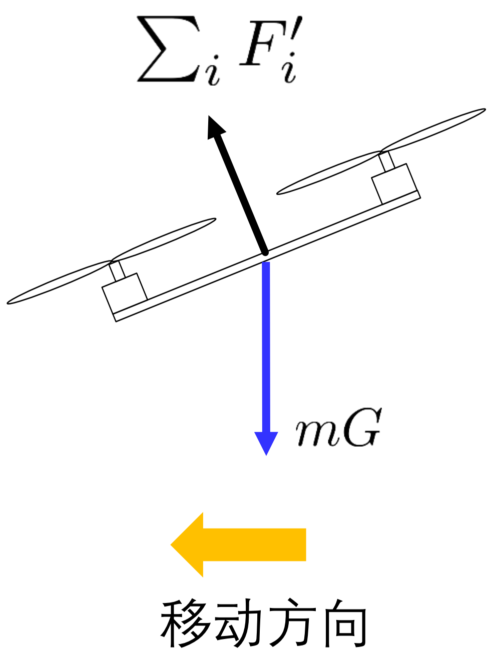
螺旋桨旋转角速度

- 因此单个螺旋桨在无人机移动情况下所产生的力：

$$F_i' = -\alpha\omega_i^2\boldsymbol{k} - \lambda\omega_i\boldsymbol{v}^{\perp}$$

旋翼无人机运动特性

- 无人机整体受力



$$m\dot{\boldsymbol{v}} = \sum_i F_i' + mG$$

$$m\dot{\boldsymbol{v}} = \sum_i(-\alpha\omega_i^2\boldsymbol{k} - \lambda\omega_i\boldsymbol{v}^{\perp}) + mG$$

世界坐标系转机体坐标系

$$\boldsymbol{v} = R\boldsymbol{v}^b$$

$$\dot{\boldsymbol{v}} = R\Omega_{\times}\boldsymbol{v}^b + R\dot{\boldsymbol{v}}^b$$

$$G = RG^b \quad (R = [\boldsymbol{i}\,\boldsymbol{j}\,\boldsymbol{k}])$$

$$\dot{\boldsymbol{v}}^b = \begin{bmatrix} G_x^b - \mu v_x - (\omega_y v_z^b - \omega_z v_y^b) \\ G_y^b - \mu v_y - (\omega_z v_x^b - \omega_z v_z^b) \\ G_z^b - a_z - (\omega_x v_y^b - \omega_y v_x^b) \end{bmatrix}$$

$$\left(a_z = \sum_i \frac{\alpha\omega_i^2}{m}, \mu = \sum_i \frac{\lambda\omega_i}{m}\right)$$

旋翼无人机运动特性

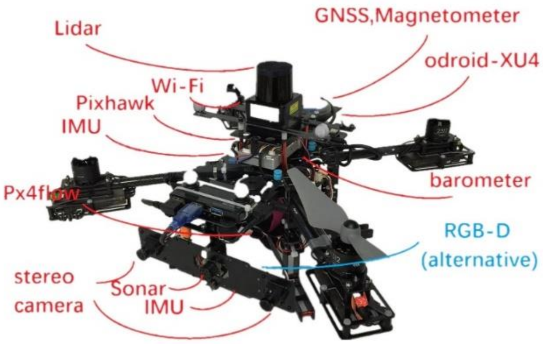
- 加速计测量比力(Special force – 注：加速计实际测量的是非引力加速度，理想状态自由落体加速计测量为零)

$$\dot{\boldsymbol{v}}^b = \begin{bmatrix} \cancel{G_x^b} - \mu v_x - (\omega_y v_z^b - \omega_z v_y^b) \\ \cancel{G_y^b} - \mu v_y - (\omega_z v_x^b - \omega_z v_z^b) \\ \cancel{G_z^b} - a_z - (\omega_x v_y^b - \omega_y v_x^b) \end{bmatrix}$$

$$\tilde{a}_x \approx -\mu v_x$$
$$\tilde{a}_y \approx -\mu v_y$$
$$\tilde{a}_z \approx -a_z$$

旋翼无人机运动特性

- 结合旋翼无人机空气动力学的状态估计方法



Wang, Rongzhi, Danping Zou, etc. "An aerodynamic model-aided state estimator for multi-rotor uavs." In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 2164-2170. IEEE, 2017.

Sartori, Daniele, Danping Zou, etc. "A Revisited Approach to Lateral Acceleration Modeling for Quadrotor UAVs State Estimation." In 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 5711-5718. IEEE, 2018.

总结

- 根据应用场景深度订制是视觉SLAM未来研究方向之一
- 深入挖掘应用场景的环境与载体特性对提升视觉SLAM的性能有很大帮助
- 根据不同应用，仍需进一步研究：
  - 几何语义SLAM
  - 复杂运动模式回归预测

謝謝