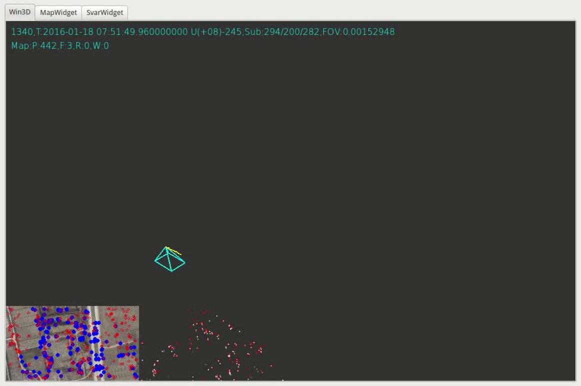
Visual SLAM and Applications
视觉SLAM及其应用

西北工业大学 - 布树辉

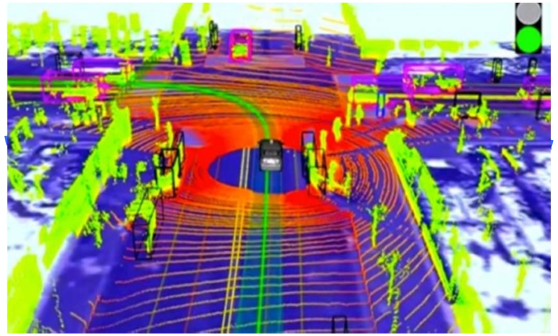
bushuhui@nwpu.edu.cn
http://www.adv-ci.com



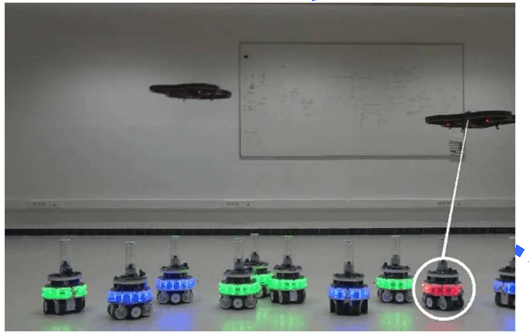
内容

- 研究背景
- 视觉SLAM
  - Map2DFusion: Real-time Mapping for UAV
  - Semi-direct Tracking and Mapping (SDTAM)
  - Place Recognition based on Deep Learning
- 展望未来

未来？ 即将到来！

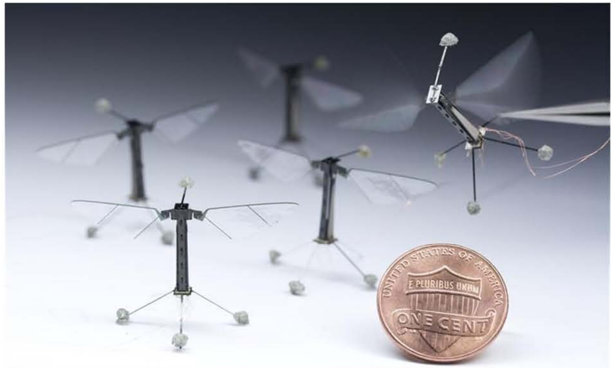


人工智能



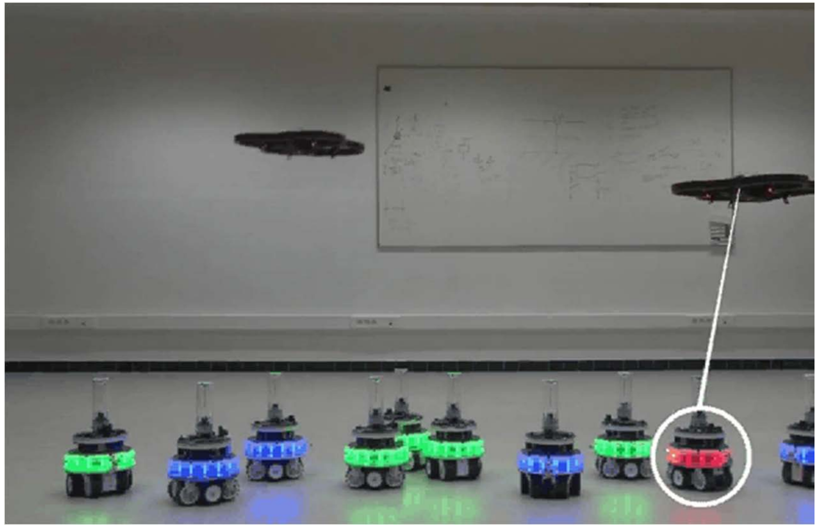
无人机 – 人工智能的载体

无人机因具有体积小、造价低、使用方便、对作战环境要求低、战场生存能力较强等优点，备受世界各国军队的青睐，广泛应用于反恐，侦查等领域。在民用领域也逐步受到极大的关注。



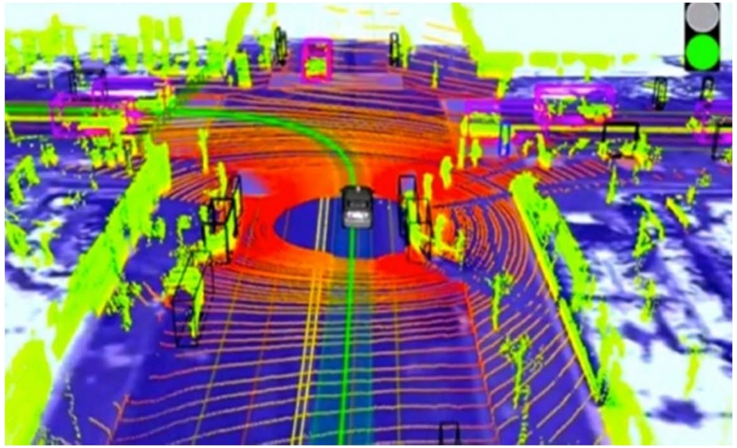
无人机 – 全自主飞行

无人机目前主要以人工操控的方式来控制其飞行或执行任务。存在较多的问题，例如工作负荷大，需要专业训练，无法控制多架飞机或编队飞机。



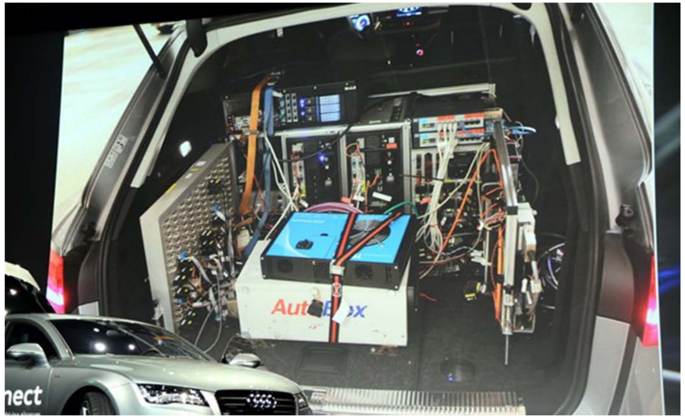
自主运行系统在其他领域

- 对于自动驾驶汽车和机器人，其未来的主要发展方向同样也是自主运行。
- 多家知名公司已基本实现汽车的自动驾驶，但其缺陷是使用较为复杂的传感器和计算设备，导致系统价格高、全天候运行能力差等。
- 目前主要的研究方向是**如何减少传感器的使用量，并提高其对环境的感知能力**，例如高精度、高实时性的行人和车辆的监测、识别等。
- 预期到**2020年左右，自动驾驶系统将成熟并实现商业化运作**。

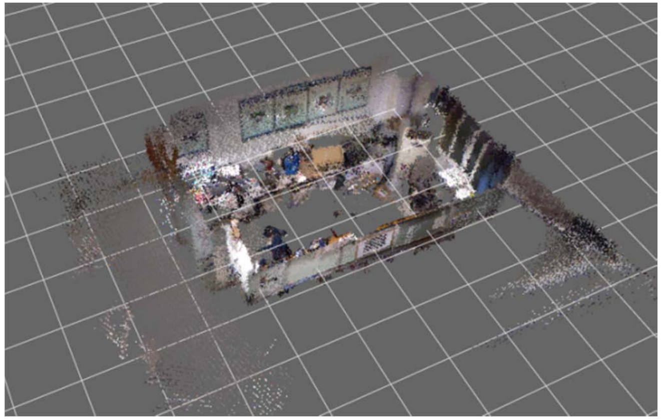


环境感知、规避障碍

汽车自动驾驶

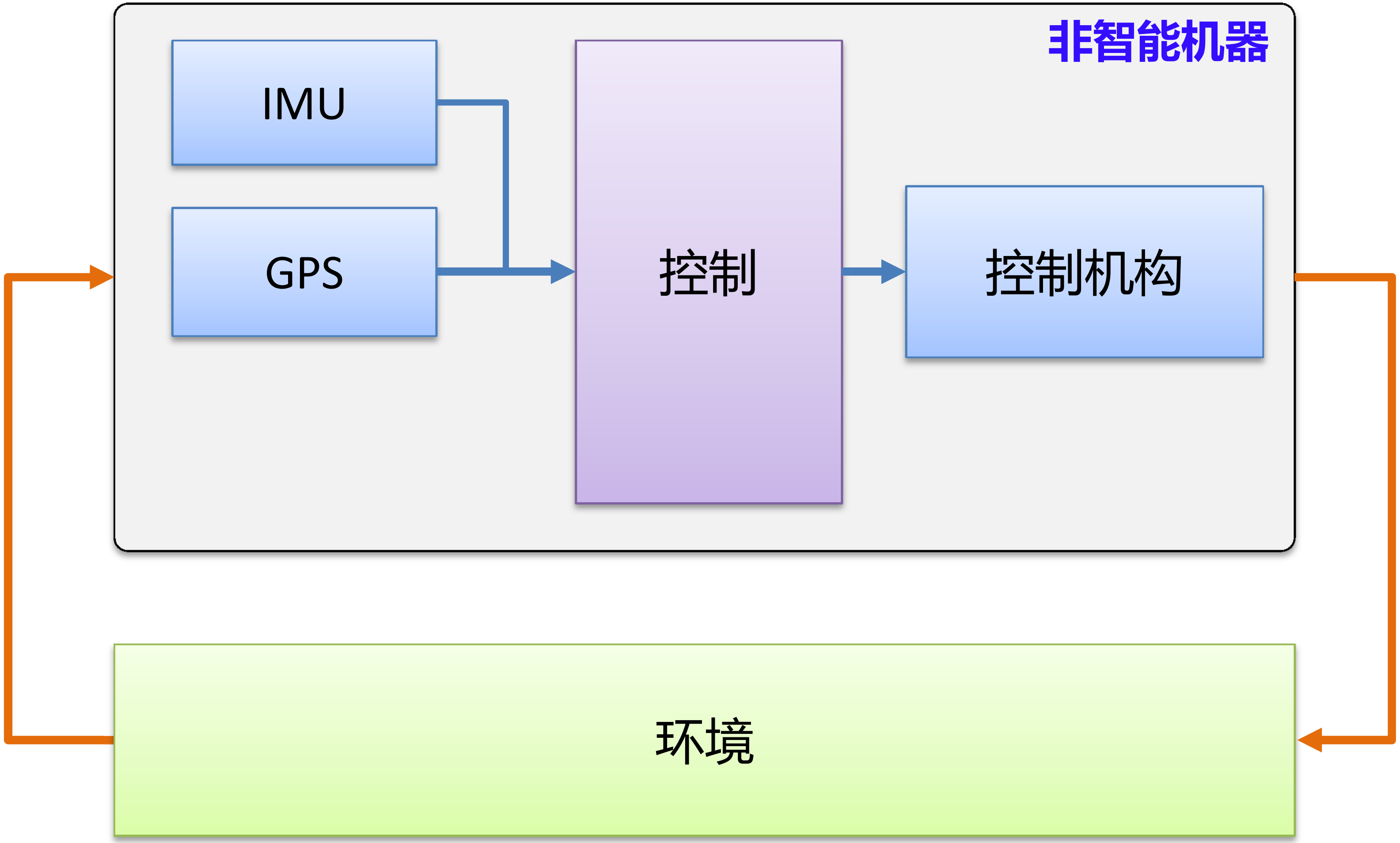


自动驾驶设备

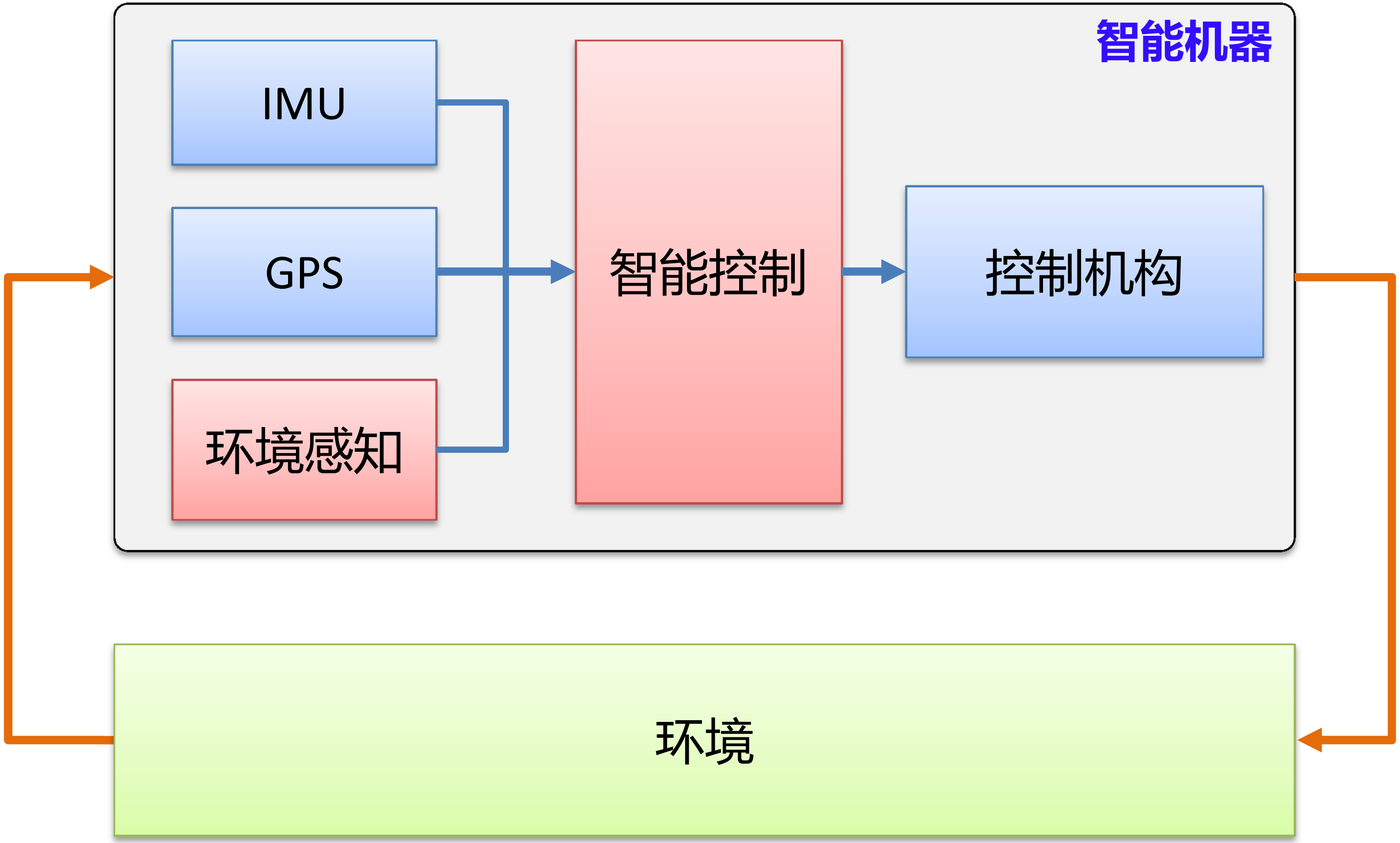


机器人室内定位、地图构建

目前成熟的



正在路上的



面临的挑战

人工智能

环境感知

推理

决策

SLAM

场景理解

多数据源
图像
LiDAR
红外
多光谱，高光谱
IMU
MAP

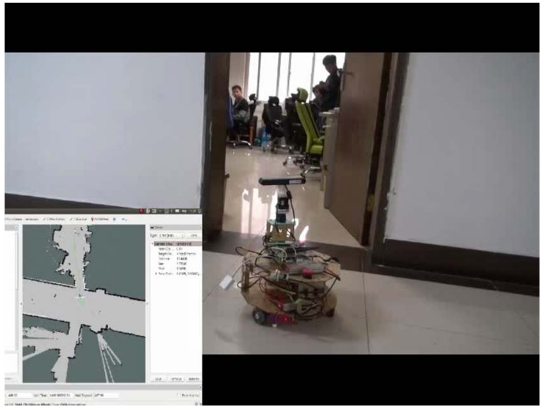
实时，低延时

复杂环境
电磁环境，气象环境

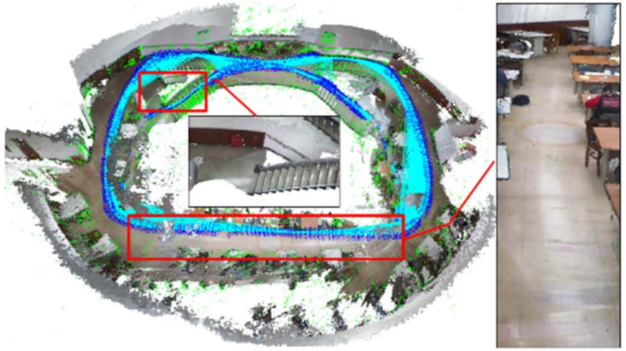
- 环境感知的效能与人相比，差距很大
- SLAM系统主要依赖几何计算、认知层面的能力较弱
- SLAM的鲁棒性、可靠性还未达到实用化程度
- 多智能体、动态环境的环境感知有待研究、开发

技术体系

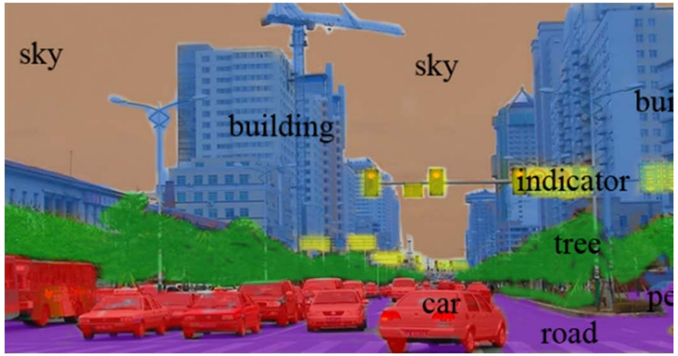
Smart Machines



Computer Vision



Machine Learning

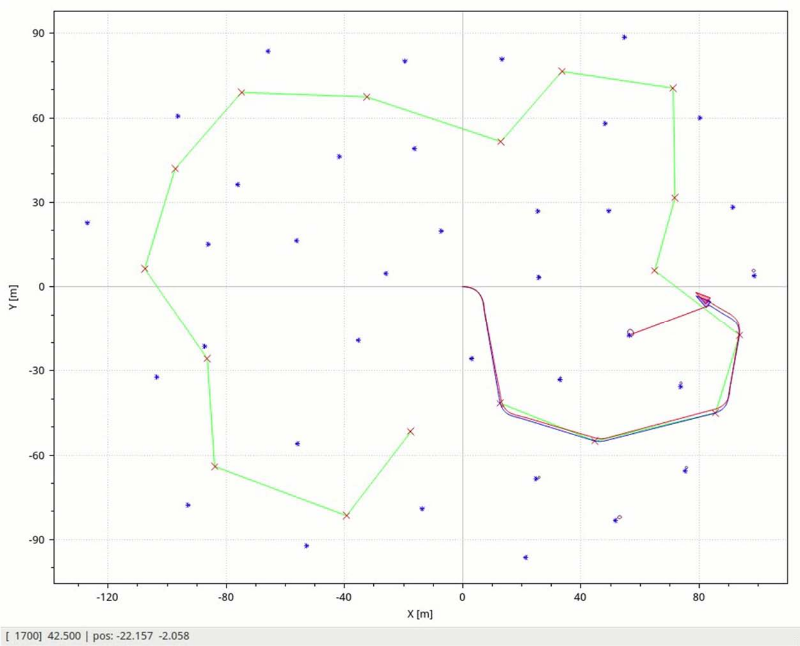


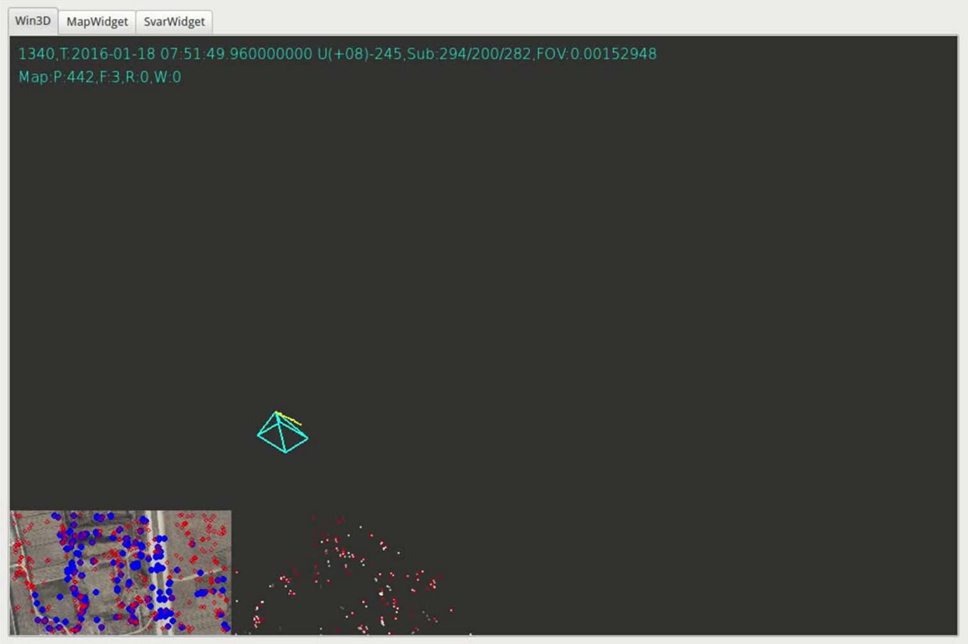
内容

- 研究背景
- 视觉SLAM
  - Map2DFusion： Real-time Mapping for UAV
  - Semi-direct Tracking and Mapping (SDTAM)
  - Place Recognition based on Deep Learning
- 展望未来

SLAM？

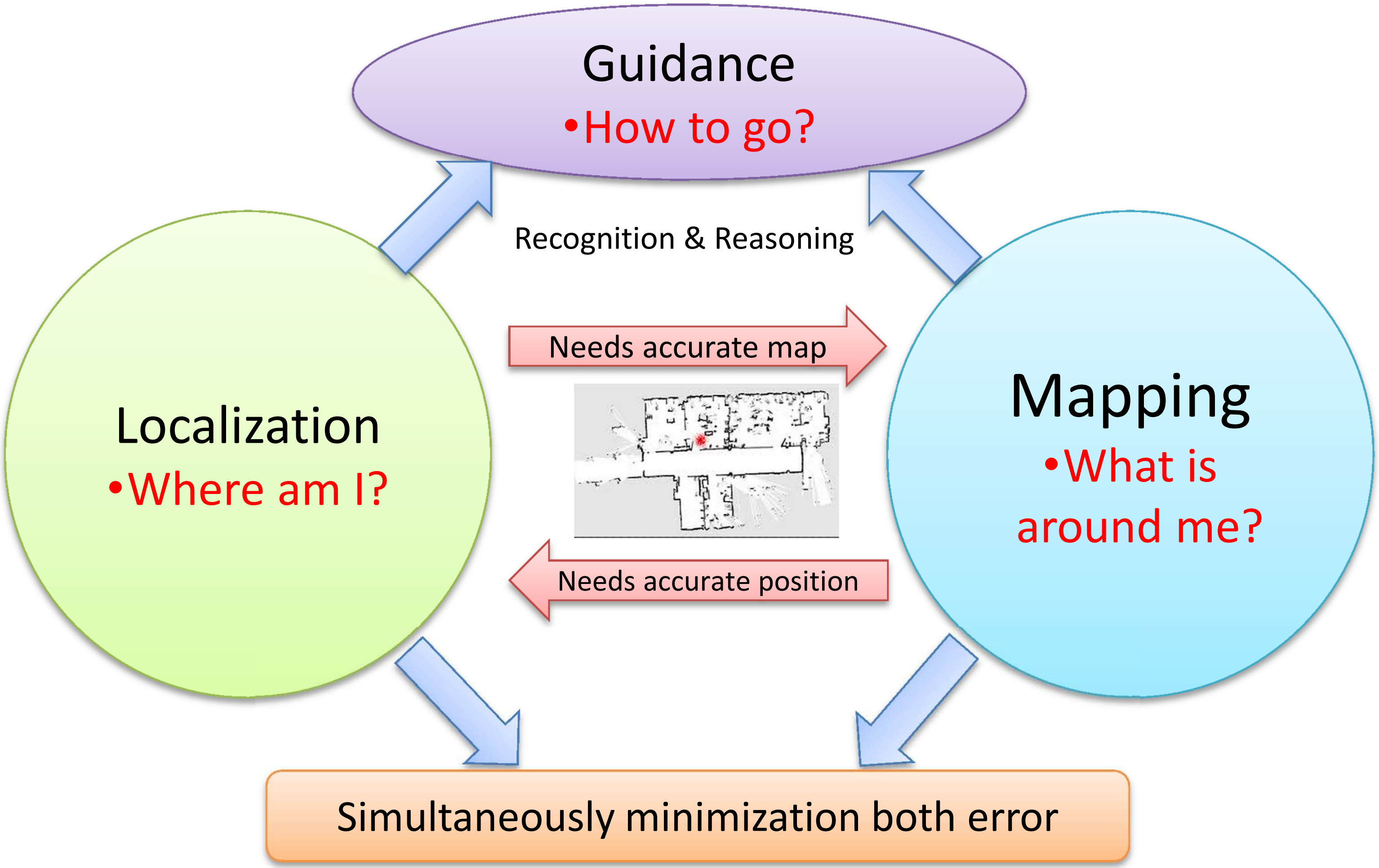
SLAM (Simultaneous Localization and Mapping)，即时定位与地图构建，最早由美国著名学者Smith于1988年提出，它是解决机器人视觉的关键核心技术，被很多学者认为是实现真正全自主移动机器人的关键。由于早期SLAM算法设计的局限性，以及传感器技术、计算机处理能力的限制，未能广泛推广应用。



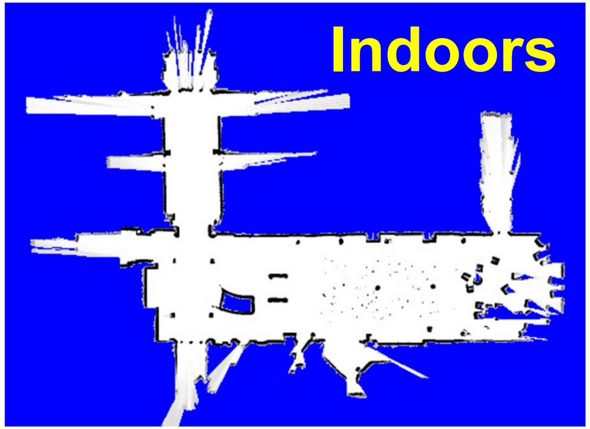


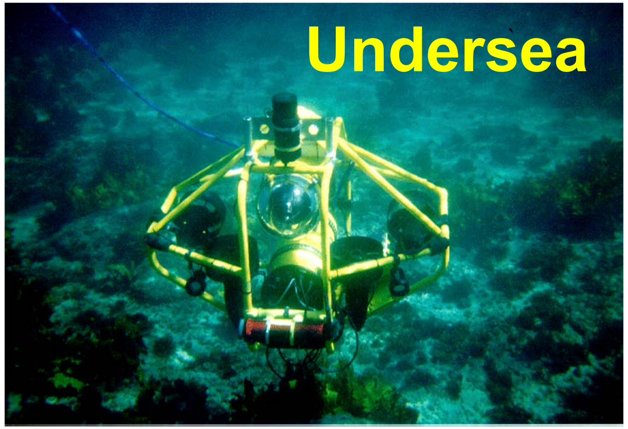
Demo can be downloaded at: http://www.adv-ci.com/blog/source/fastslam-gui/

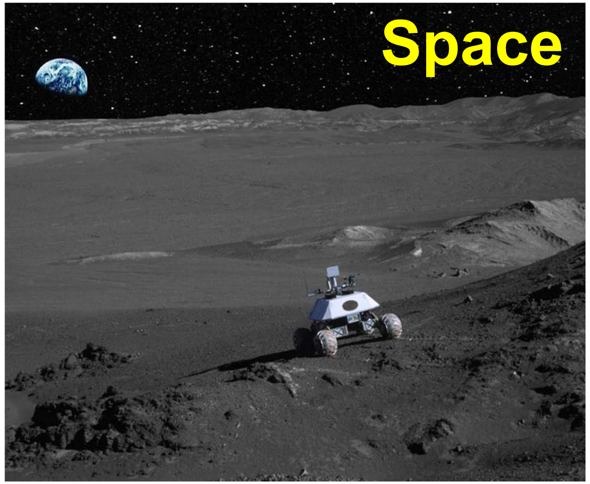
SLAM – 导航与环境感知



SLAM – 广阔的应用领域

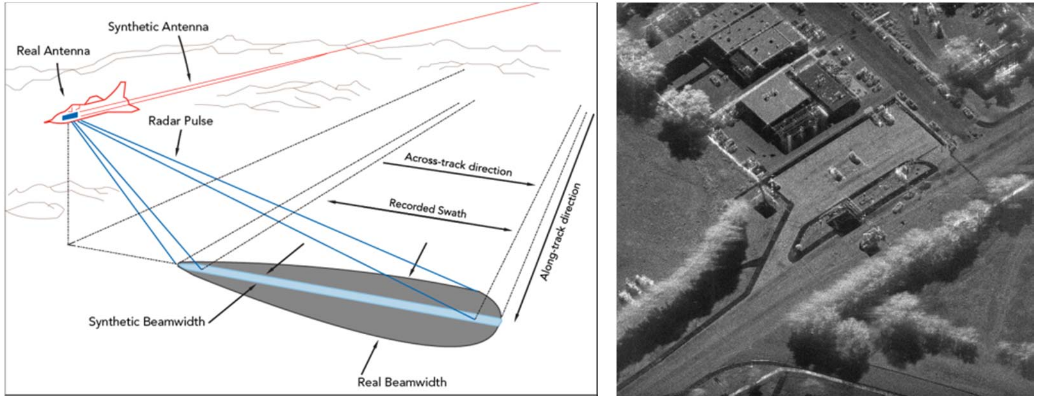




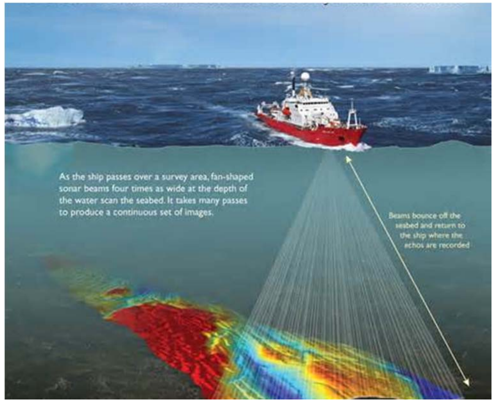


AR

SLAM – 传感器



SAR



SONAR

LiDAR

Camera

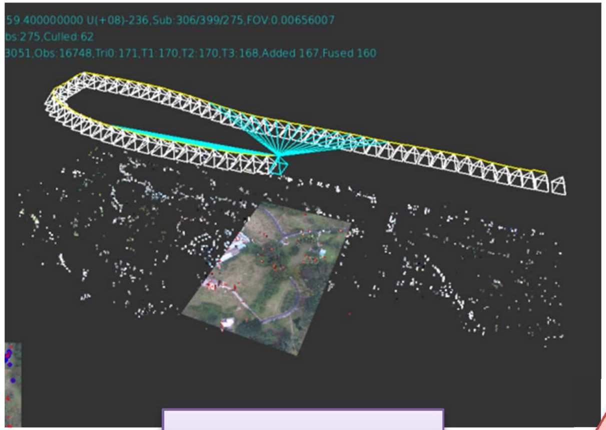
SLAM – 视觉? 激光?

视觉	激光
低成本	高成本
cm ~ km 范围	cm ~ m 范围
丰富的信息	只有距离信息
精度相对低	精度高
与人类似的感知	超人类的感知
视觉SLAM难度高（门槛高）	激光SLAM难度低（门槛低）
应用前景广泛	应用范围受限

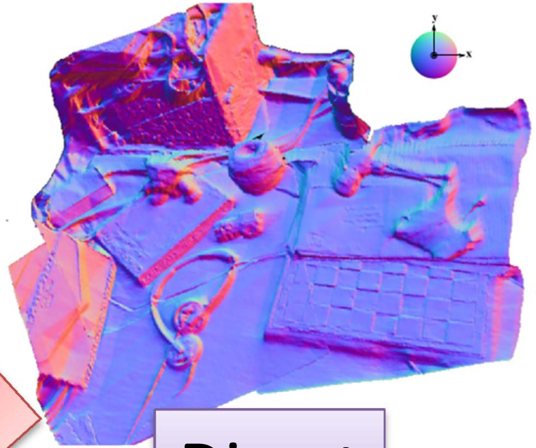
视觉

激光

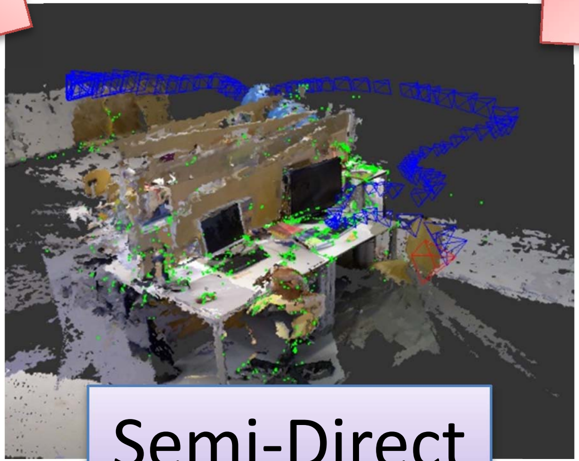
视觉SLAM – 基本分类



Feature

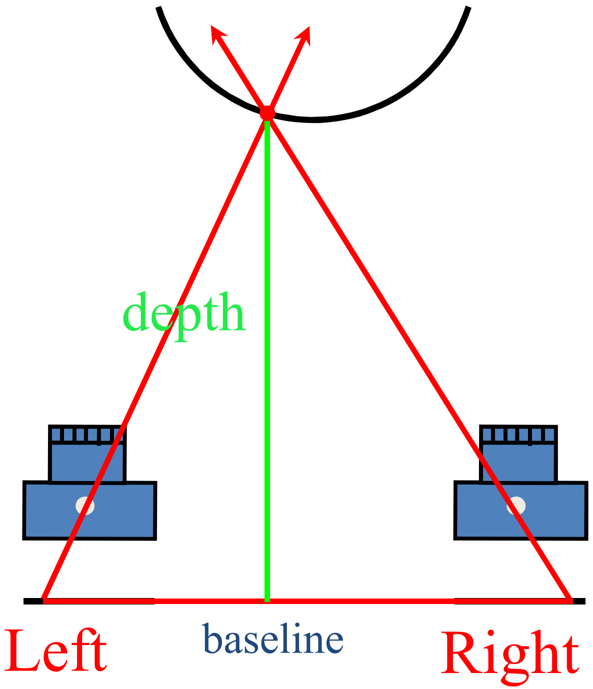


Direct



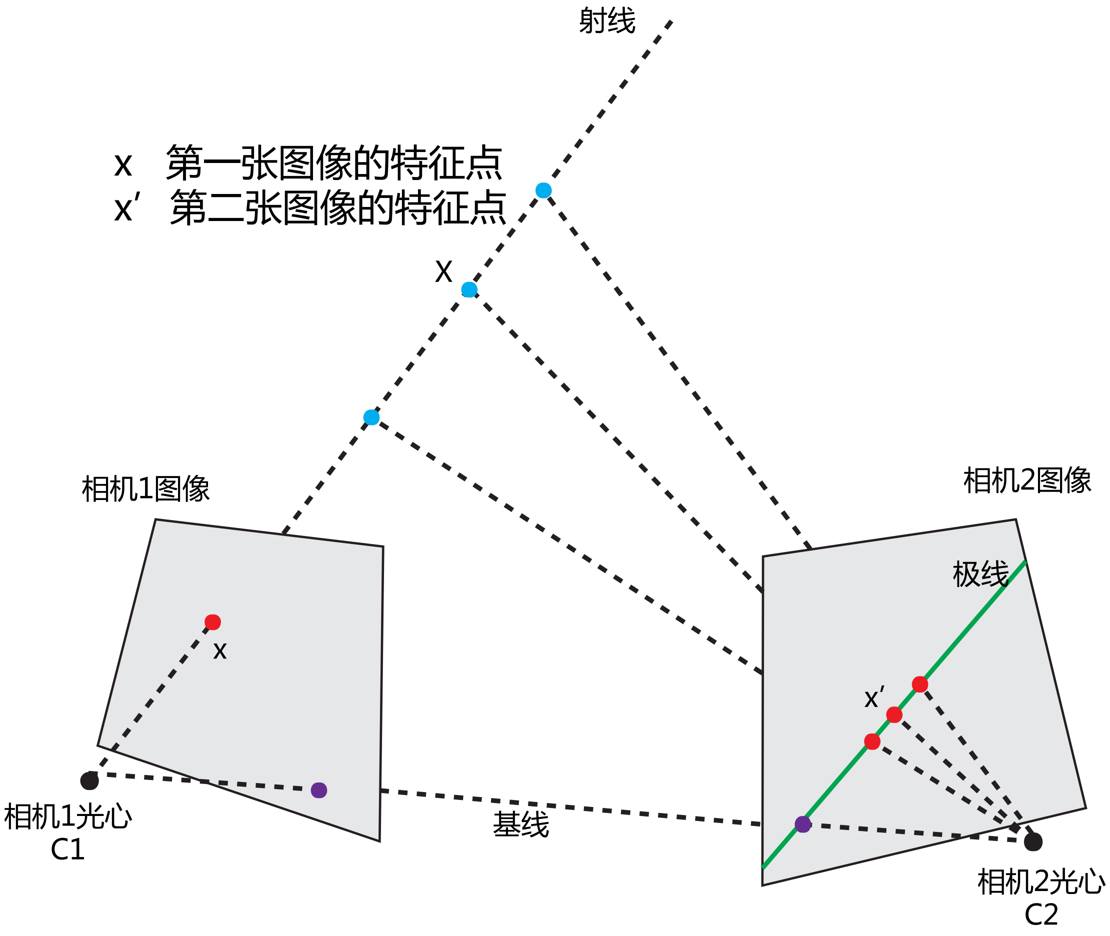
Semi-Direct

视觉SLAM原理 – 基本原理



通过三角化两张图像上对应的点，能够恢复物体点到相机的相对深度值

视觉SLAM原理 – 如何计算相机的位置和姿态



$$\mathrm{x}'^{\mathrm{T}} F x = 0$$

几组特征对应关系得到 F

↓ F，基本矩阵

$$\mathrm{E} = \mathrm{K}^{\mathrm{T}}\mathrm{F}\mathrm{K}$$

F矩阵和相机内参矩阵得到本征矩阵 E

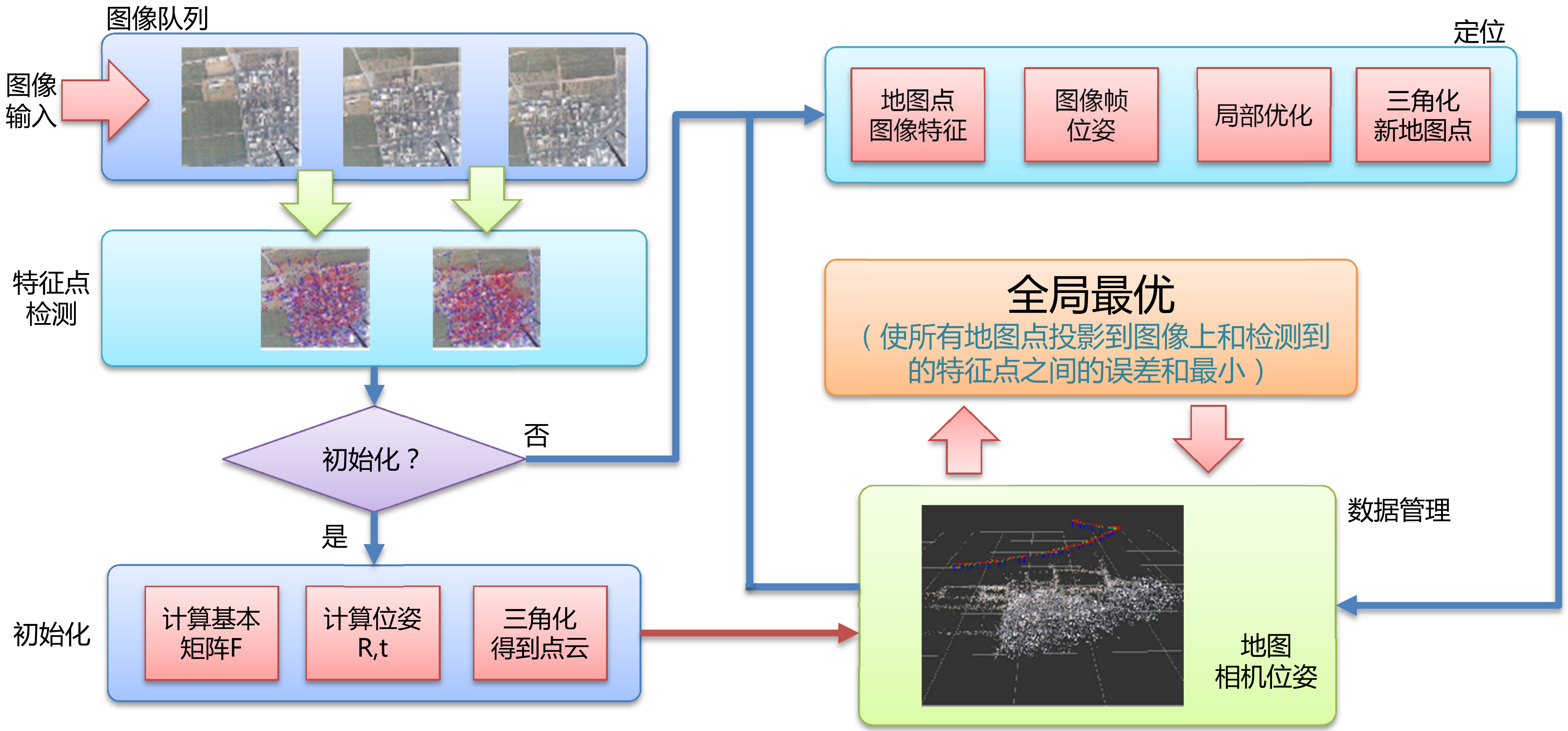
↓ E

R, t

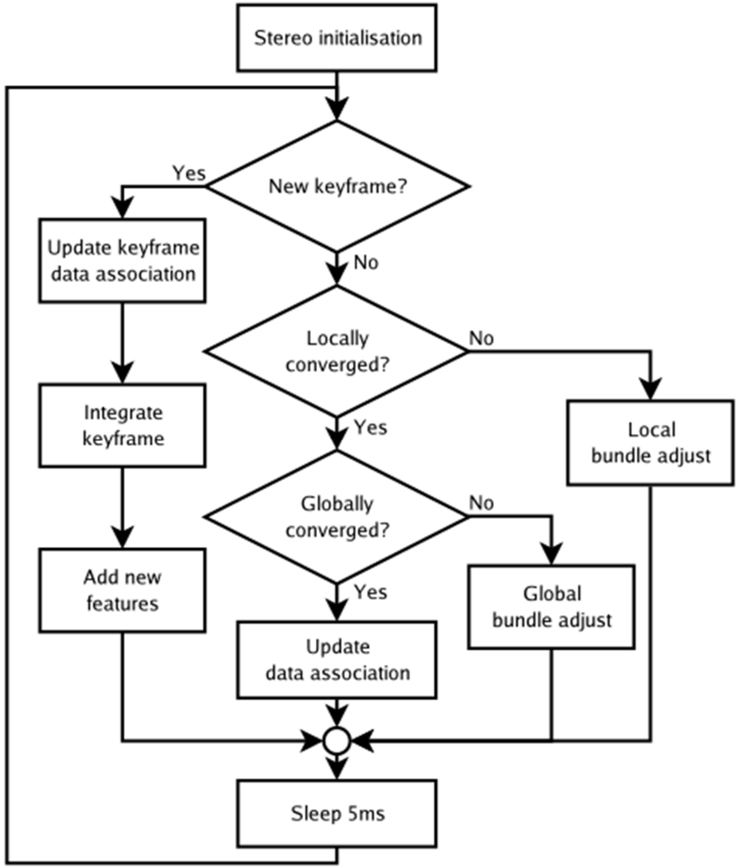
得到第二张图像的位置，姿态

假设第一张图像位于远点，可以得到第二张图像的位置和姿态

视觉SLAM原理 – Feature Based Methods

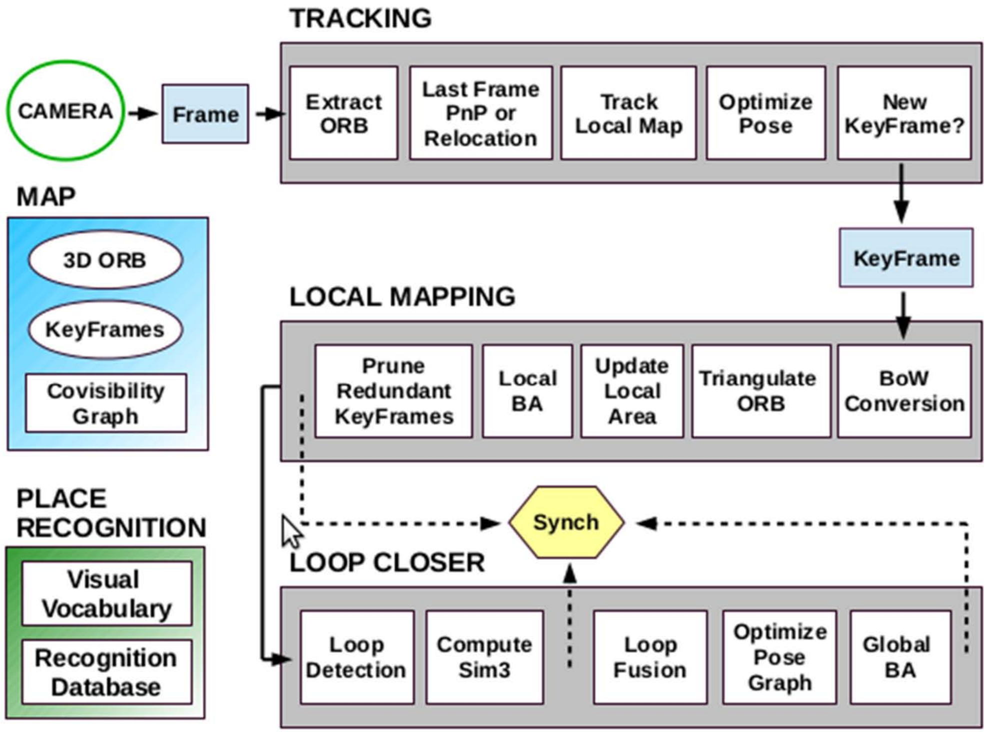


Feature Based Methods - Parallel Tracking and Mapping



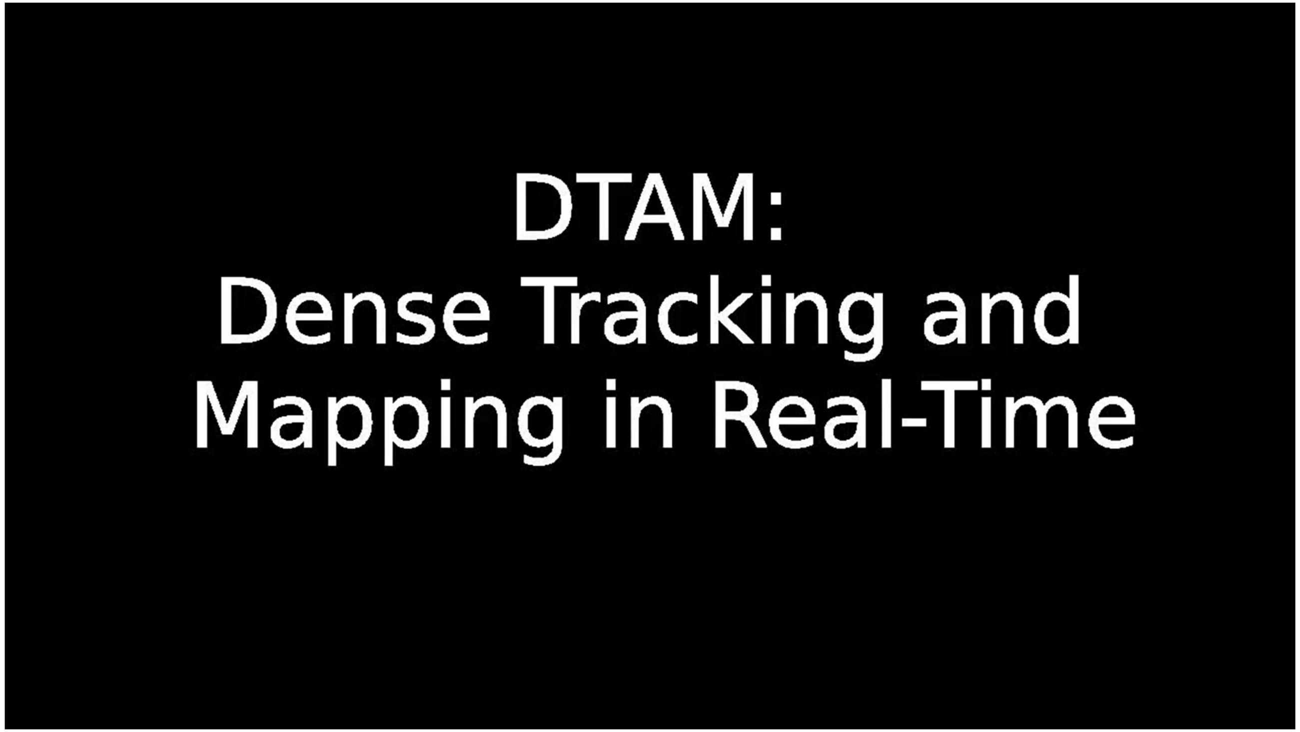
G. Klein, D. Murray, Parallel Tracking and Mapping for Small AR Workspaces, 2007

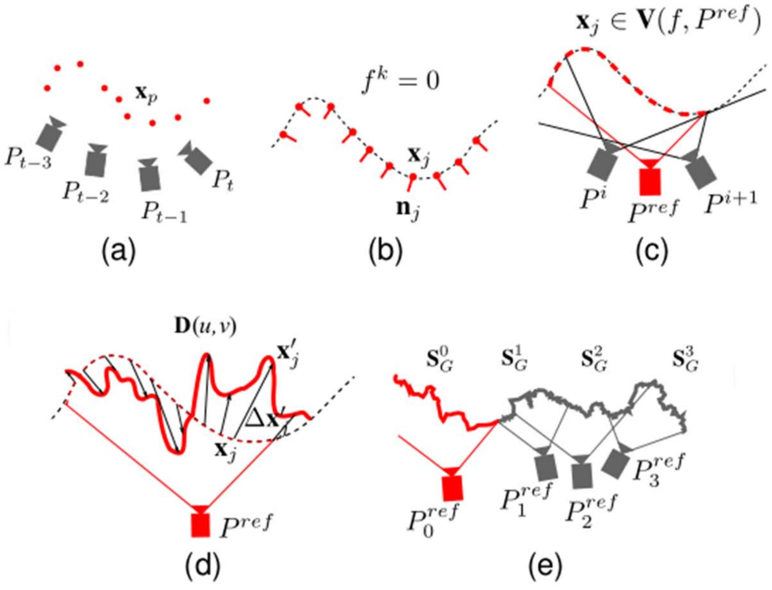
Feature Based Methods – ORB-SLAM



Raúl Mur-Artal, J. M. M. Montiel and Juan D. Tardós. ORB-SLAM: A Versatile and Accurate Monocular SLAM System, 2015

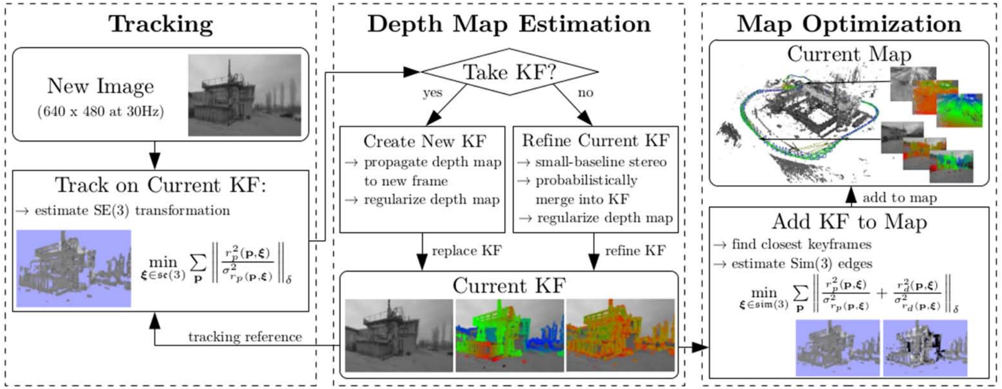
Direct Based Methods – Dense Tracking and Mapping (DTAM)





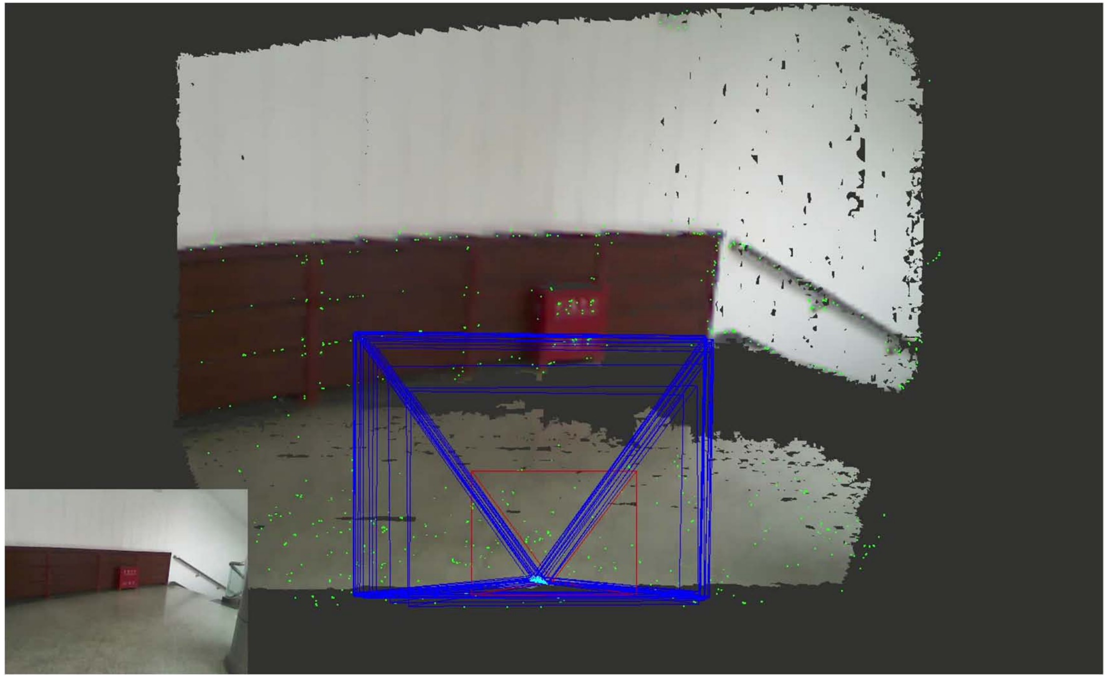
R.A. Newcombe, S.J. Lovegrove and A.J. Davison, DTAM: Dense Tracking and Mapping in Real-Time, 2011

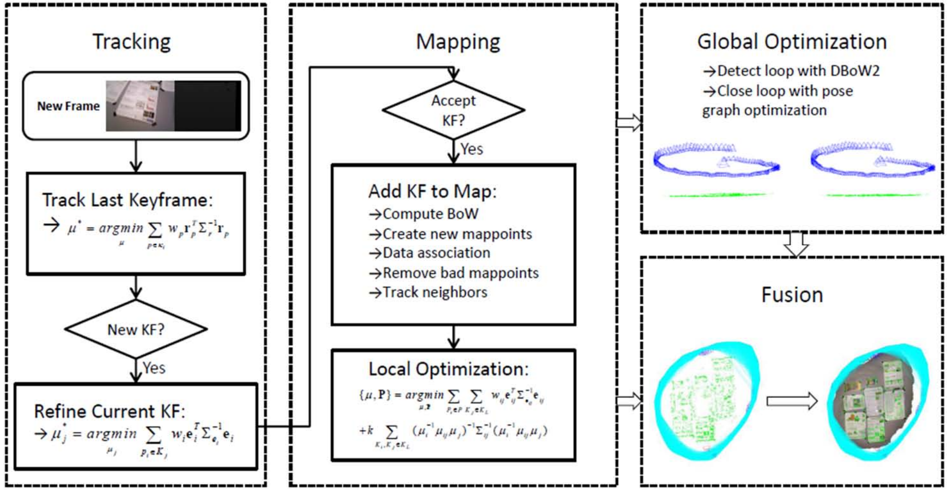
Direct Based Methods – Large-Scale Direct Monocular SLAM (LSD-SLAM)



J. Engel, T. Schops, and D. Cremers, LSD-SLAM: Large-Scale Direct Monocular SLAM, 2014

Semi-direct Tracking and Mapping (SDTAM)





- The direct method is adopted to track current motion with high-speed, followed with a motion refinement based on feature correspondences.
- Balance between accuracy and efficiency can be realized.

Shuhui Bu, Yong Zhao, et al., Semi-direct Tracking and Mapping with RGB-D Camera for MAV, MTAP, 2016.

内容

- 研究背景
- 视觉SLAM
  - Map2DFusion： Real-time Mapping for UAV
  - Semi-direct Tracking and Mapping (SDTAM)
  - Place Recognition based on Deep Learning
- 展望未来

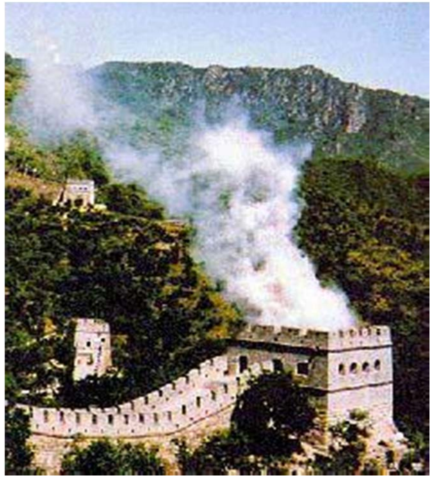
战争与地图

兵法有云：

“知己知彼，百战不殆”

- 凡是战争都离不开信息，现代战争是信息主导的战争，信息在战争中具有十分显著的地位和作用。
- 作战地图是作战指挥人员策划战术的重要依据，现阶段往往通过遥感、航拍甚至实地勘测等手段制作高精度的军事地图。

在实际作战中战场环境却不是一成不变的。

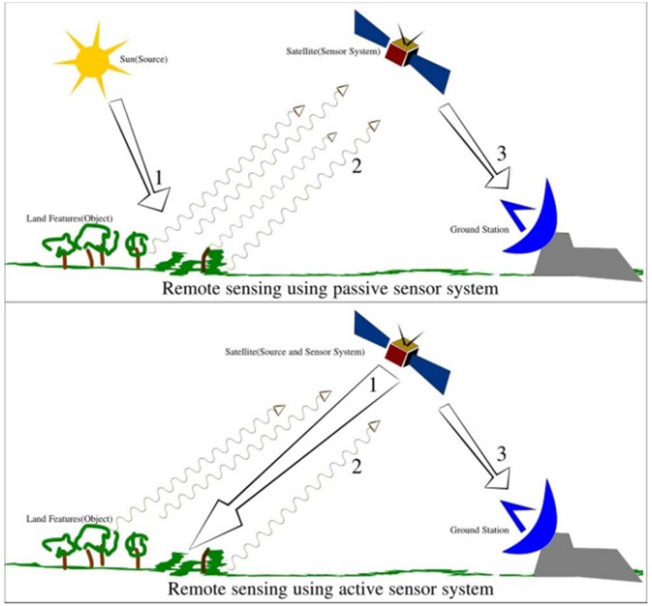


烽火

三维电子作战地图

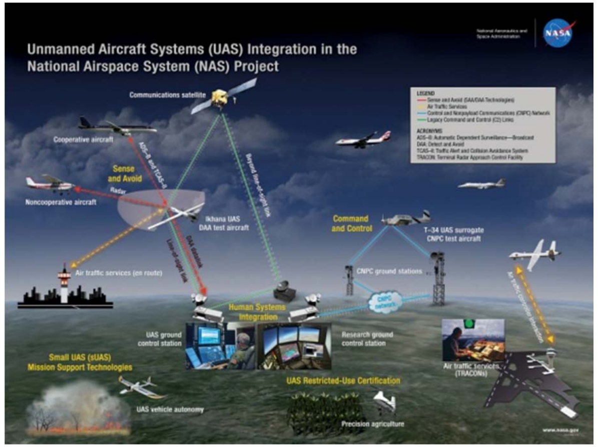
传统方法面临的问题

一般采用遥感的方式，是一种非接触的，远距离的探测技术，运用传感器对物体的电磁波的辐射、反射特性的探测。面临的主要问题是实时性较弱，无法即时得到环境地图。



无人机自主飞行

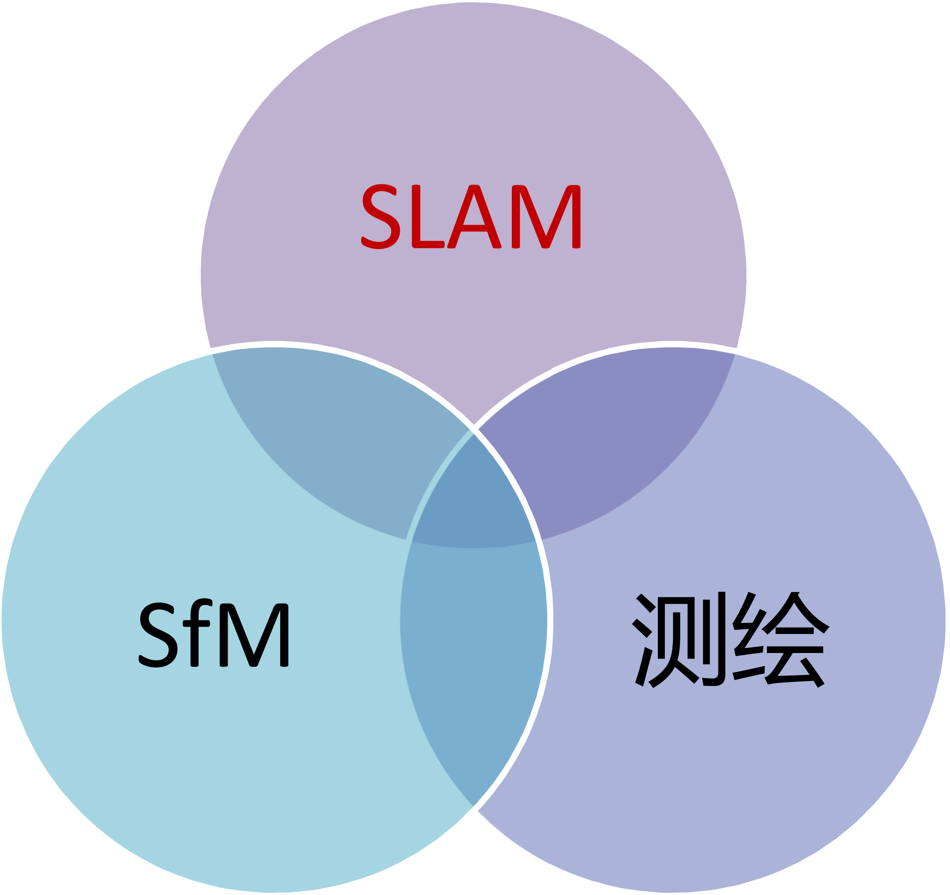
由于环境的复杂，无人机数据链路中断，GPS受到干扰等情况下，必须自主飞行。自主飞行需要对环境进行感知，从而指导无人机的飞行。在无GPS的情况下，需要多种导航手段来提供导航信息，实时地图是一种重要的导航信息源。



解决之道

SLAM与SfM对比

- SLAM更多考虑图像之间的关系
- SLAM对算法做了更多的优化，能够达到实时
- **设计巧妙的算法能够实时融合采集的图像得到DOM**



SLAM与测绘方法对比

- 能够达到90%以上的精度
- 使用POS数据的方式不同，SLAM能够进行联合优化
- SLAM对算法做了更多的优化，能够达到实时
- **定位与地图生成同步进行，能够用到机器人自主导航**

SfM即Structure from Motion，计算机视觉的三维重建方法，Pix4D等使用的方法

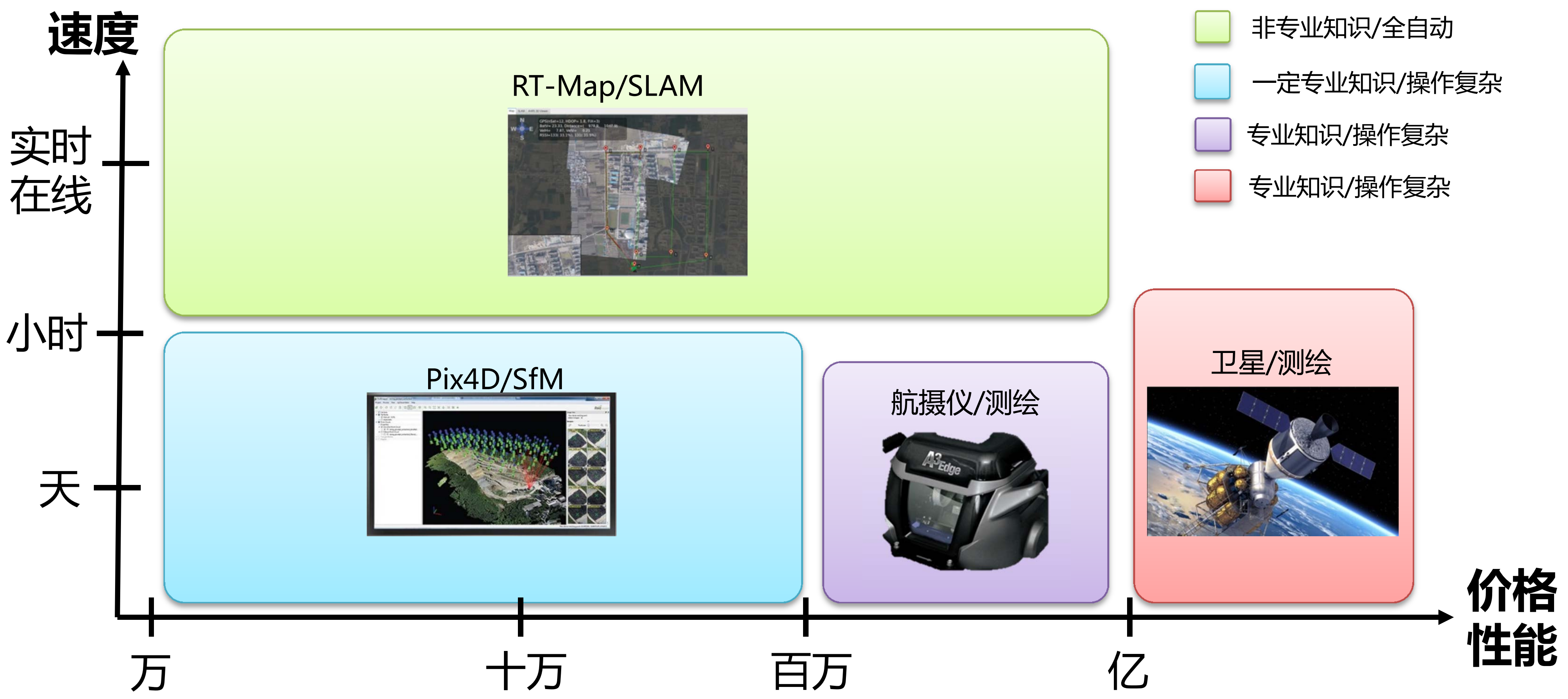
解决方案

摇杆控制

- 高可靠性
- 操作简单
- 自动飞行

- 实时数据处理
- G-SLAM
- 鲁棒性高

弥补传统方法的实时性不足



需求与解决方法

Requirements:

1. It must be able to achieve real-time speed.
2. The stitching needs to be incremental for both blending and visualization.
3. Camera contains not only rotation but also translation.
4. The mosaicing results should be as ortho as possible and able to be updated properly.

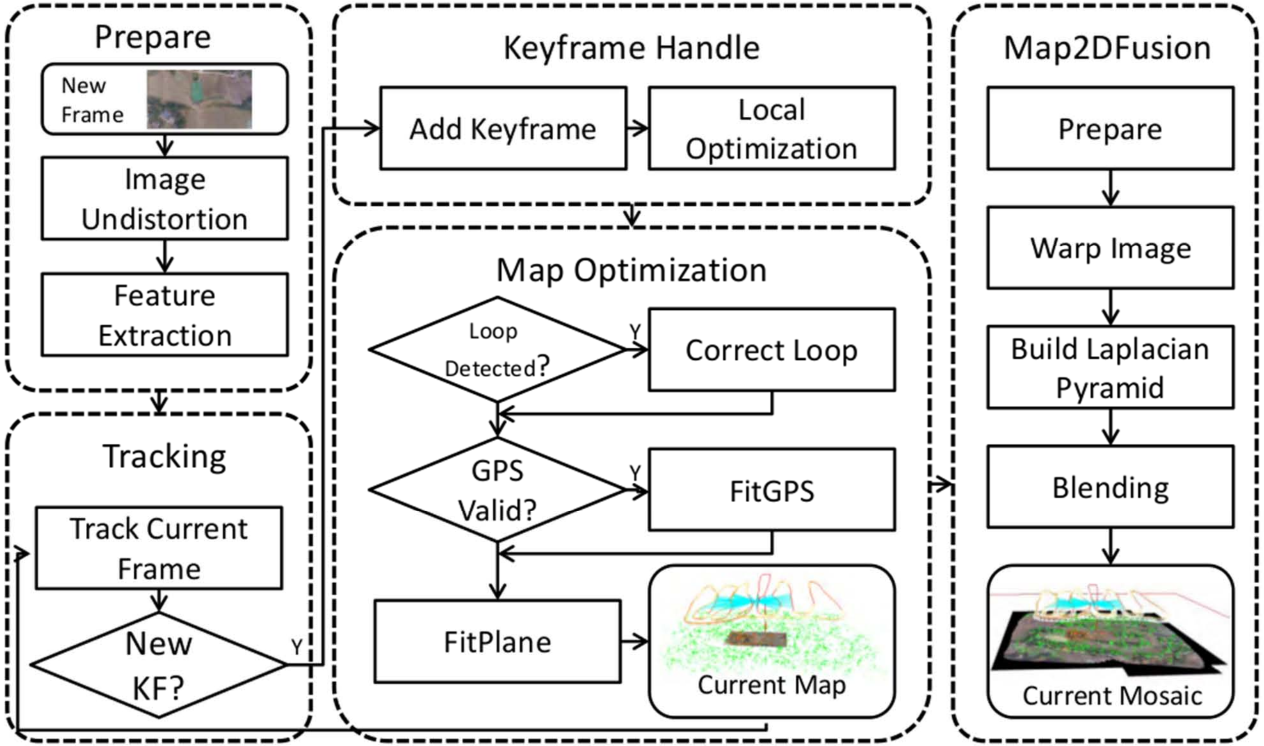
Adopt SLAM to estimate camera's attitude and position

The reconstructed point cloud is used to estimate the terrain

Using a novel adaptive weighted multi-band blender to stitching images

Shuhui Bu, Yong Zhao, et al., Map2DFusion: Real-time Incremental UAV Image Mosaicing based on Monocular SLAM, IROS, 2016

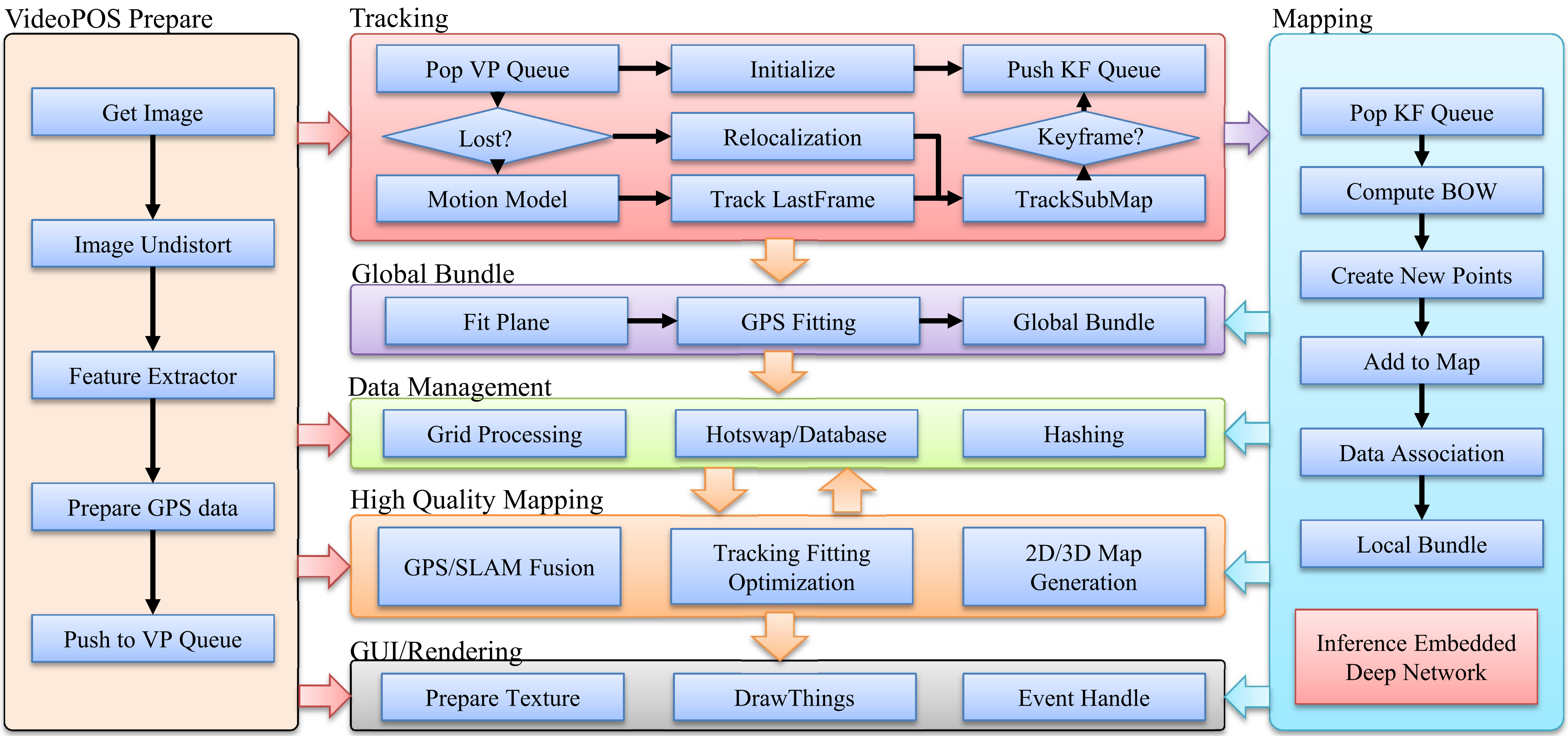
软件系统 – 核心算法



- **Feature based Visual SLAM System: G-SLAM**
- **Automatic GPS and video synchronization**: a graph based optimization is proposed to synchronize video time with GPS time from coarse to fine.
- **Real-time orthoimage blender**: an adaptive weighted multi-band method to blend and visualize images incrementally in real-time.

Shuhui Bu, Yong Zhao, et al., Map2DFusion: Real-time Incremental UAV Image Mosaicing based on Monocular SLAM, IROS, 2016

特点1：G-SLAM系统架构



特点2：Automatic GPS and video synchronization

$$\{\mu^*, \mathbf{P}^*\} = \underset{\mu,\mathbf{P}}{argmin} \sum_{P_i \in P_L} \sum_{K_j \in K_L} w_{ij} \mathbf{e}_{ij}^T \Sigma_{ij}^{-1} \mathbf{e}_{ij} + \sum_{K_j \in K_L} \gamma(j)^T \Lambda_j^{-1} \gamma(j),$$

Reprojection error

GPS alignment error

$$\gamma(j) = \mathbf{P}_{t_g}^{gps} - \mathbf{S}_{vg} \mathbf{P}_{t_j}^{map} = \mathbf{P}_{t_j + t_{vg}}^{gps} - \mathbf{S}_{vg} \mathbf{P}_{t_j}^{map}$$

$$\{t_{vg}^*, \mathbf{S}_{vg}^*\} = \underset{t_{vg}, \mathbf{S}_{vg}}{argmin} \sum_{K_j \in \mathbf{K}} \gamma(j)^T \Lambda_j^{-1} \gamma(j),$$

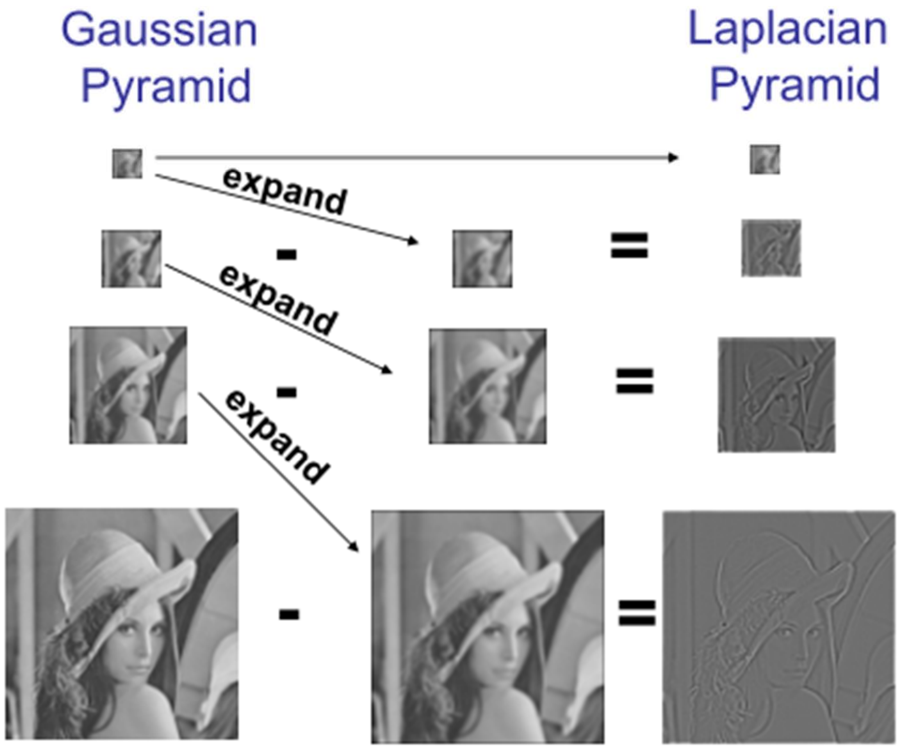
t_{vg} Time difference between GPS and image

S_{vg} SIM3 transformation from local to WGS84

Shuhui Bu, Yong Zhao, et al., Map2DFusion: Real-time Incremental UAV Image Mosaicing based on Monocular SLAM, IROS, 2016

特点3： Real-time Ortho Image Blender

Multiband Pyramid



$$G_l(x,y) = \sum_{d_x=-n}^{n} \sum_{d_y=-n}^{n} w_{d_x,d_y} G_{l-1}(x+d_x, y+d_y),$$

$$L_l = G_l - G_{l+1} \qquad l < k.$$

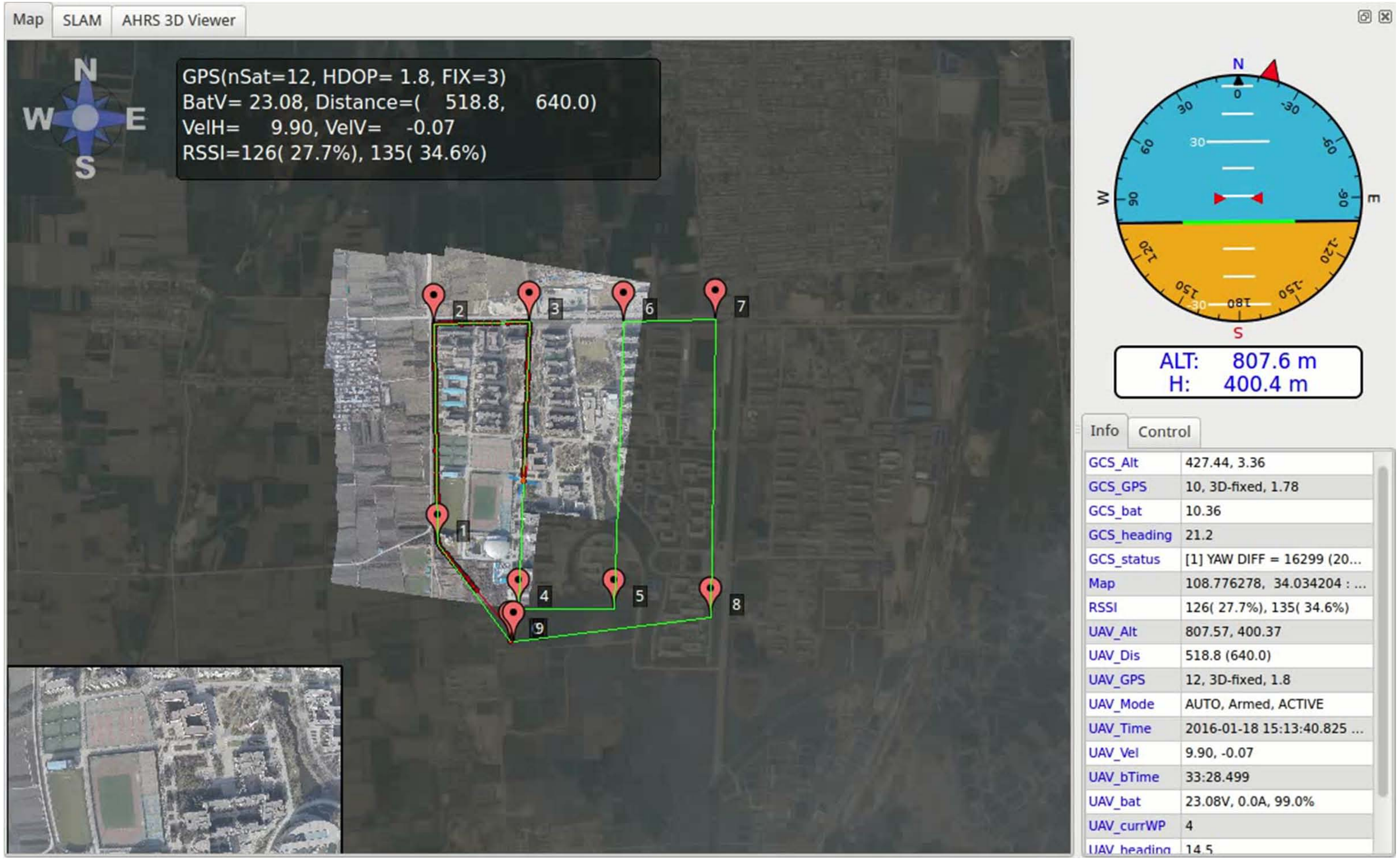
G_l Gaussian pyramid

L_l Laplacian pyramid

w_{d_x,d_y} Adaptive weight considering the height, view angle, and pixel location

Shuhui Bu, Yong Zhao, et al., Map2DFusion: Real-time Incremental UAV Image Mosaicing based on Monocular SLAM, IROS, 2016

效果展示 – 实时地图

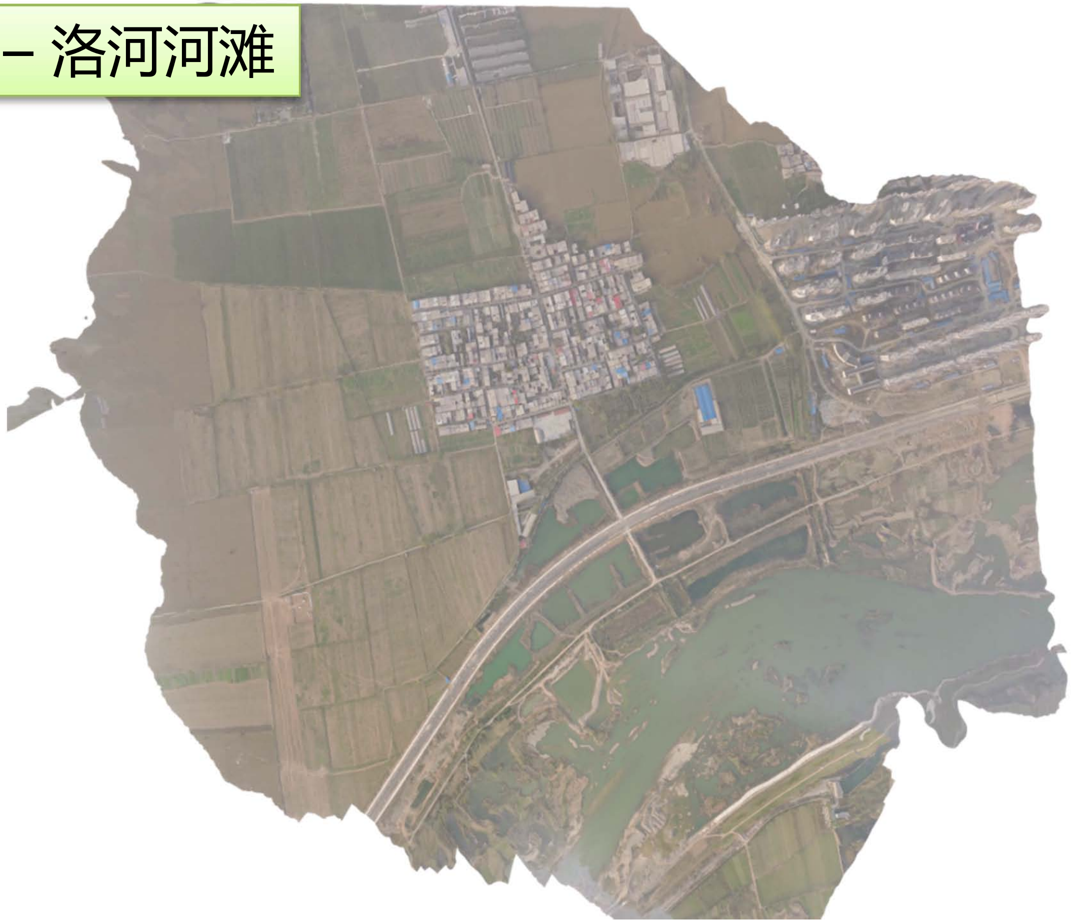


Map
SLAM
AHRS 3D Viewer
N
W
E
S
GPS(nSat=12, HDOP= 1.8, FIX=3)
BatV= 23.08, Distance=(518.8, 640.0)
VelH= 9.90, VelV= -0.07
RSSI=126(27.7%), 135(34.6%)
1
2
3
4
5
6
7
8
9
ALT: 807.6 m
H: 400.4 m
Info
Control
GCS_Alt 427.44, 3.36
GCS_GPS 10, 3D-fixed, 1.78
GCS_bat 10.36
GCS_heading 21.2
GCS_status [1] YAW DIFF = 16299 (20...
Map 108.776278, 34.034204 : ...
RSSI 126(27.7%), 135(34.6%)
UAV_Alt 807.57, 400.37
UAV_Dis 518.8 (640.0)
UAV_GPS 12, 3D-fixed, 1.8
UAV_Mode AUTO, Armed, ACTIVE
UAV_Time 2016-01-18 15:13:40.825 ...
UAV_Vel 9.90, -0.07
UAV_bTime 33:28.499
UAV_bat 23.08V, 0.0A, 99.0%
UAV_currWP 4
UAV_heading 14.5

效果展示 – 卫星地图对比

洛阳 – 洛河河滩

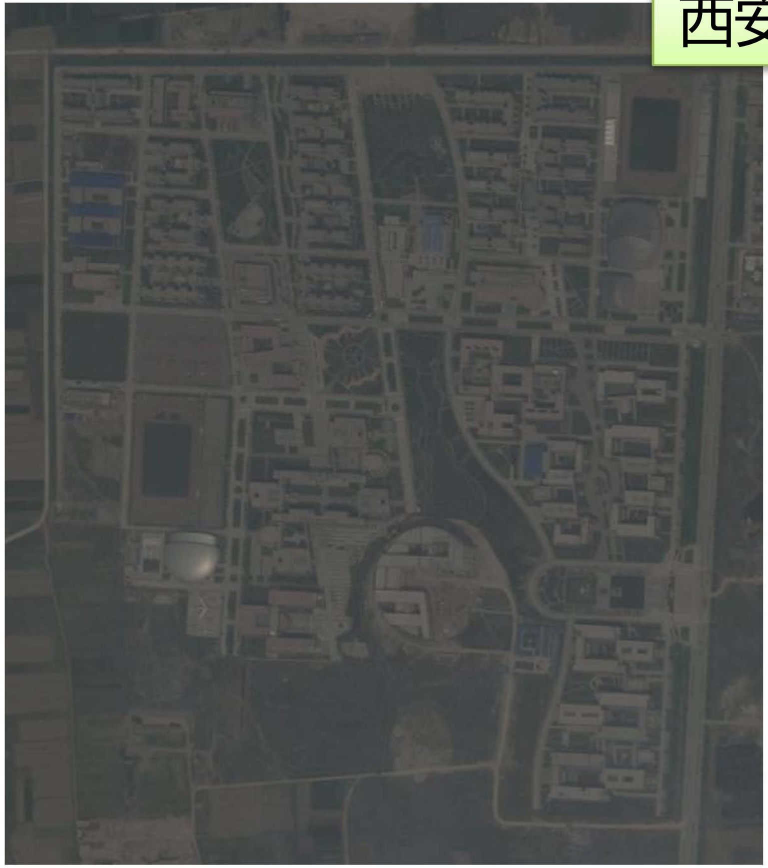
卫星地图



RT-Map

效果展示 – 卫星地图对比

西安 – 西北工业大学



卫星地图

RT-Map

效果对比　一样？其实大不同！

PhotoScan

Pix4D

RT-Map

效果对比 1

PhotoScan

Pix4D

RT-Map

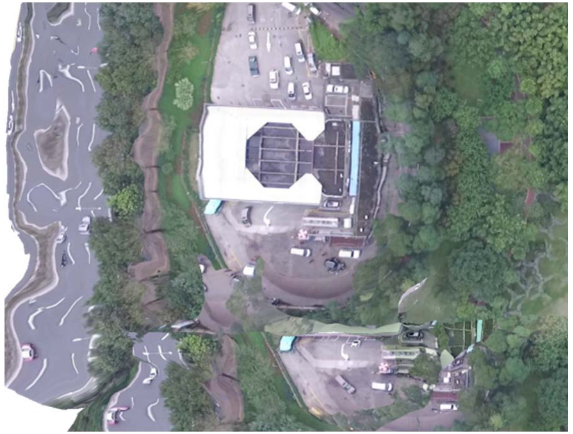
效果对比 2

PhotoScan

Pix4D

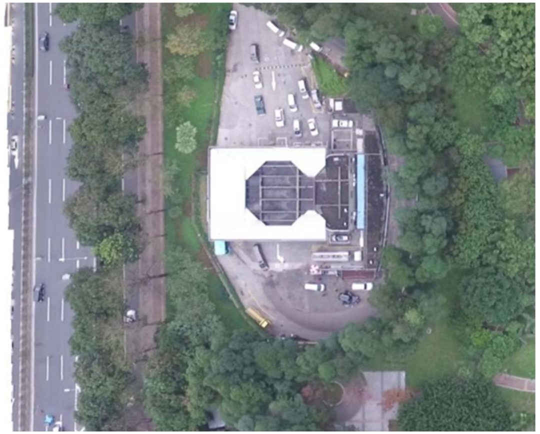
RT-Map

效果对比 3



PhotoScan

Pix4D

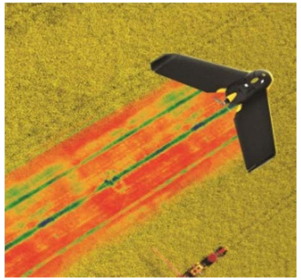


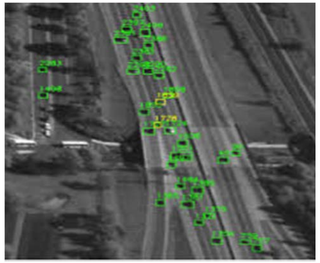
RT-Map

相关解决方法的对比

	Pix4D PhotoScan	DroneDeploy	航空摄影 测绘 倾斜摄影测量	本系统 RT-Map
实时性	离线计算（数小时）	网络云计算（需要联网）	离线计算（数天）	在线实时
安全性	一般	较低	一般	高
测量精度	中	中	高	中
多信息源融合	无	无	具备	具备
地面站整合能力	无	无	无	具备
系统硬件要求	低	较低	高	较低
二次开发能力	无	无	无	具备
使用成本	高	高	很高	低
功能	DEM，DOM，三维重建	DEM，DOM，三维重建	DEM，DOM，三维重建	DEM，DOM，三维重建 地图构建、无GPS导航定位

应用领域





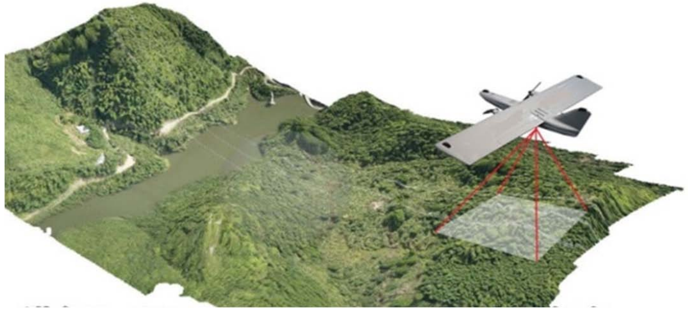
民用

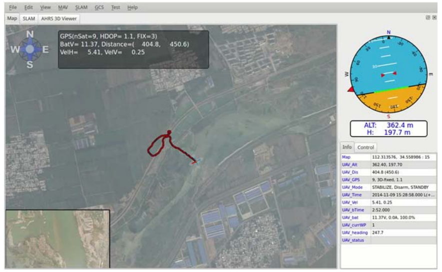
军用

其他

抗震救灾（灾害评估）
农业植保（病虫害评估）
自动驾驶，智慧城市

特定区域持续监视
侦查，无人机作战
集群作战，过饱和攻击



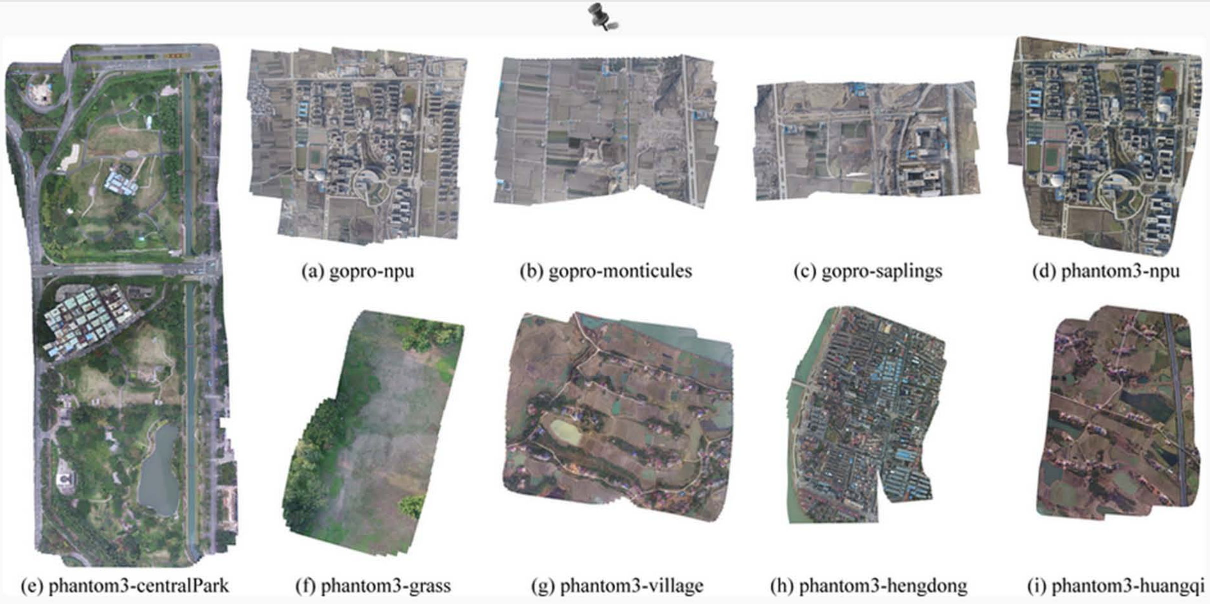


NPU Drone-Map Dataset & Sourcecode

Source code: https://github.com/zdzhaoyong/Map2DFusion
Dataset: http://www.adv-ci.com/blog/source/npu-drone-map-dataset/

NPU Drone-Map Dataset

This dataset contains several aerial video sequences captured in different terrains and heights, which is used to evaluate the effectiveness of our proposed open-source Map2DFusion system. Until now, the dataset is consisted of several sequences recorded in different locations with over 100GB data and more captured sequences will be added. Welcome to supply us more original flight data and share them on this website!



NPU DroneMap Dataset

内容

- 研究背景
- 视觉SLAM
  - Map2DFusion： Real-time Mapping for UAV
  - Semi-direct Tracking and Mapping (SDTAM)
  - Place Recognition based on Deep Learning
- 展望未来

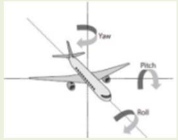
目标

Color & Depth Images

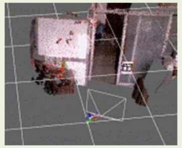
Flight Control
Path Planning
Obstacle Avoidance

Align color and depth image

Estimate position and attitude



Select keyframe and update map



Point Cloud

Position

Flight Path

实验设置

Kinect

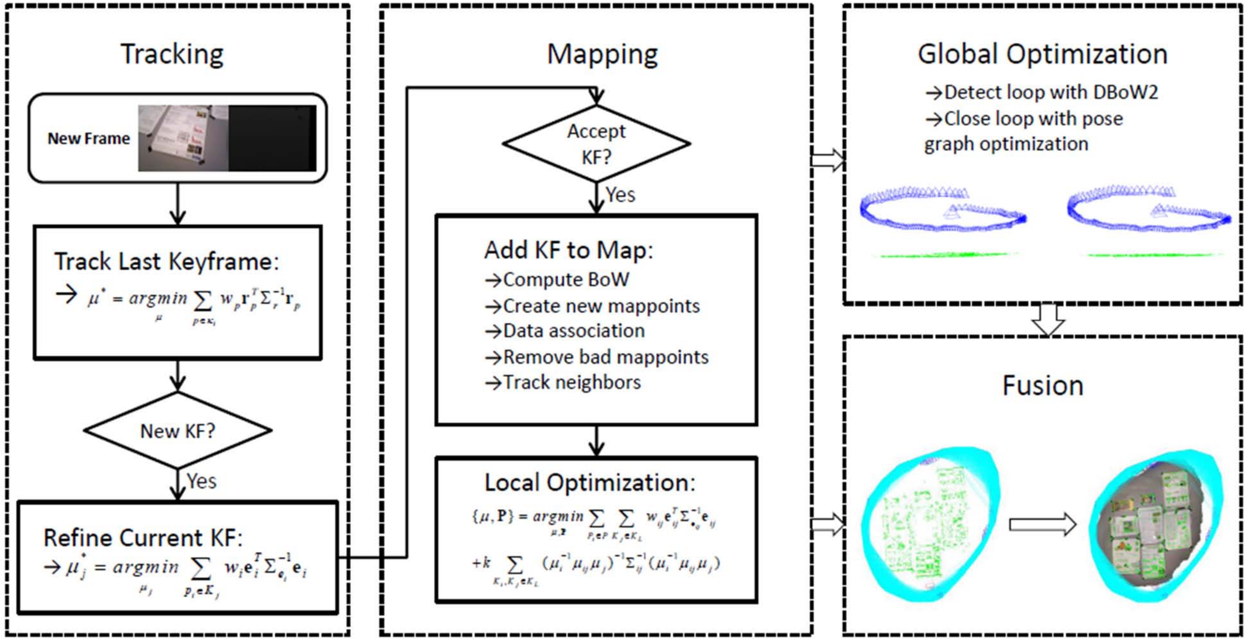
Capture color and depth images

ArDrone 2

Indoor flight plateform

Experiment setup

Semi-direct Tracking and Mapping



Feature:

- Direct method is adopted to track current motion with high-speed, followed with a motion refinement based on feature correspondences.
- A novel error function based on mixed depth and geometric measurements is designed to achieve high accuracy and robustness for pose estimation.

Shuhui Bu, Yong Zhao, et al., Semi-direct Tracking and Mapping with RGB-D Camera for MAV, MTAP, 2016.

Direct Tracking

$$p' = warp(\mu_{ji}, p) = Proj(P') = Proj(exp(\mu_{ji}) \cdot Proj^{-1}(p)).$$

Point in keyframe K_j

SE3 transformation from i->j

Point in keyframe K_i

$$r_p = \begin{bmatrix} I_i(p) - I_j(p') \\ [p']_d - [P']z \end{bmatrix},$$

Photometric and depth errors

$$\mu_{ji}^{*} = \underset{\mu}{argmin} \sum_{p \in K_i} w_p r_p^T \Sigma_r^{-1} r_p,$$

The optimized relative pose

$$w_p = \frac{v+1}{v + r_p^T \Sigma^{-1} r_p}.$$

Weight based on t-distribution $p_t(0, \Sigma, v)$

Shuhui Bu, Yong Zhao, et al., Semi-direct Tracking and Mapping with RGB-D Camera for MAV, MTAP, 2016.

Feature based Mapping – Pose Refinement

Detect and Extract ORB feature points

Optimize current pose u_j through matching keypoints between K_j and K_i

Create local sub-map which composed of keyframes share same observations

Local optimization

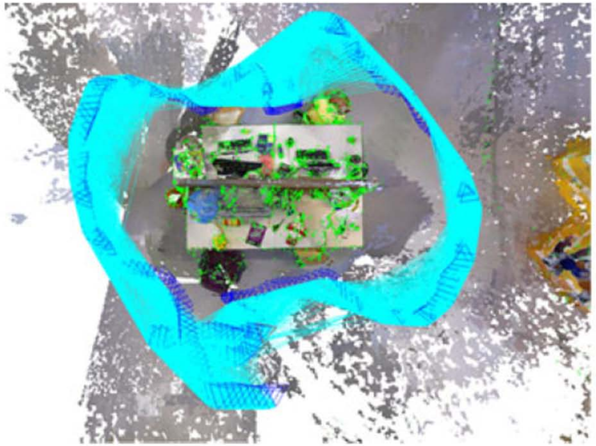
$$\mathbf{e} = \begin{bmatrix} e_x \\ e_y \\ e_d \end{bmatrix} = \begin{bmatrix} x \\ y \\ 1/d \end{bmatrix} - Proj(\exp(\mu_j)\mathbf{P})$$

$$\mu_j^* = \underset{\mu_j}{argmin} \sum_{p_i \in K_j} w_i \mathbf{e}_i^T \Sigma_{e_i}^{-1} \mathbf{e}_i.$$

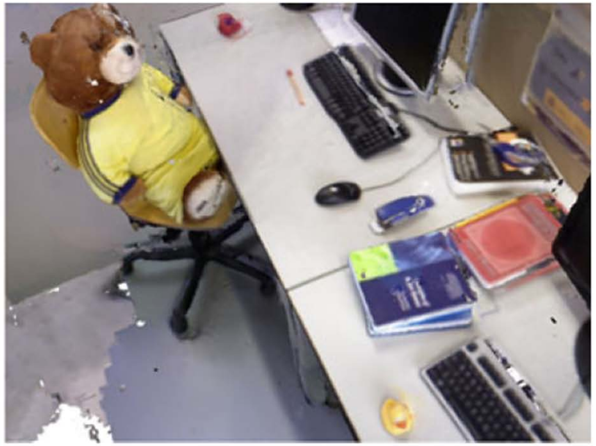
$$w_i = \begin{cases} \left(1 - \frac{\mathbf{r}_i^T \Sigma_{e_i}^{-1} \mathbf{r}_i}{\sigma}\right)^2, & \mathbf{r}_i^T \Sigma_{e_i}^{-1} \mathbf{r}_i \le \sigma, \\ 0, & \mathbf{r}_i^T \Sigma_{e_i}^{-1} \mathbf{r}_i > \sigma. \end{cases}$$

$$\{\mu, \mathbf{P}\} = \underset{\mu, \mathbf{P}}{argmin} \sum_{P_i \in P} \sum_{K_j \in K_L} w_{ij} \mathbf{e}_{ij}^T \Sigma_{\mathbf{e}_{ij}}^{-1} \mathbf{e}_{ij}$$
$$+ k \sum_{K_i, K_j \in K_L} (\mu_i^{-1} \mu_{ij} \mu_j)^{-1} \Sigma_{ij}^{-1} (\mu_i^{-1} \mu_{ij} \mu_j),$$

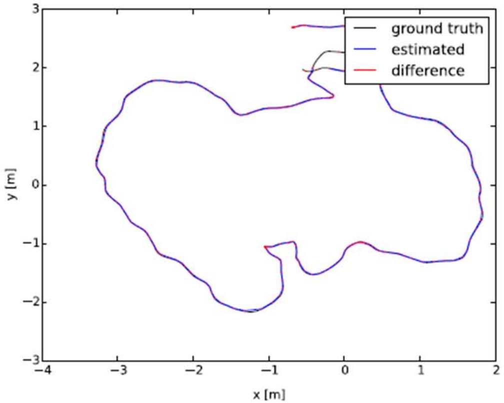
Experiment Results (TUM Dataset)



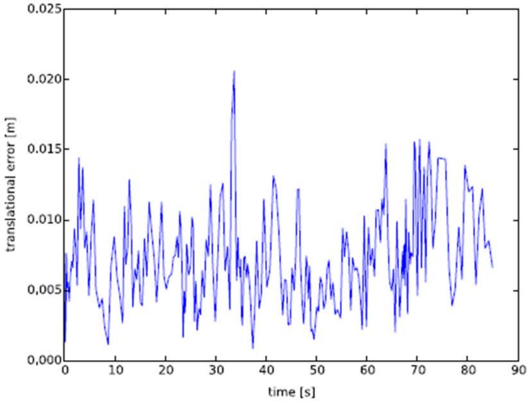
(a) Reconstruction result



(b) Fusion detail



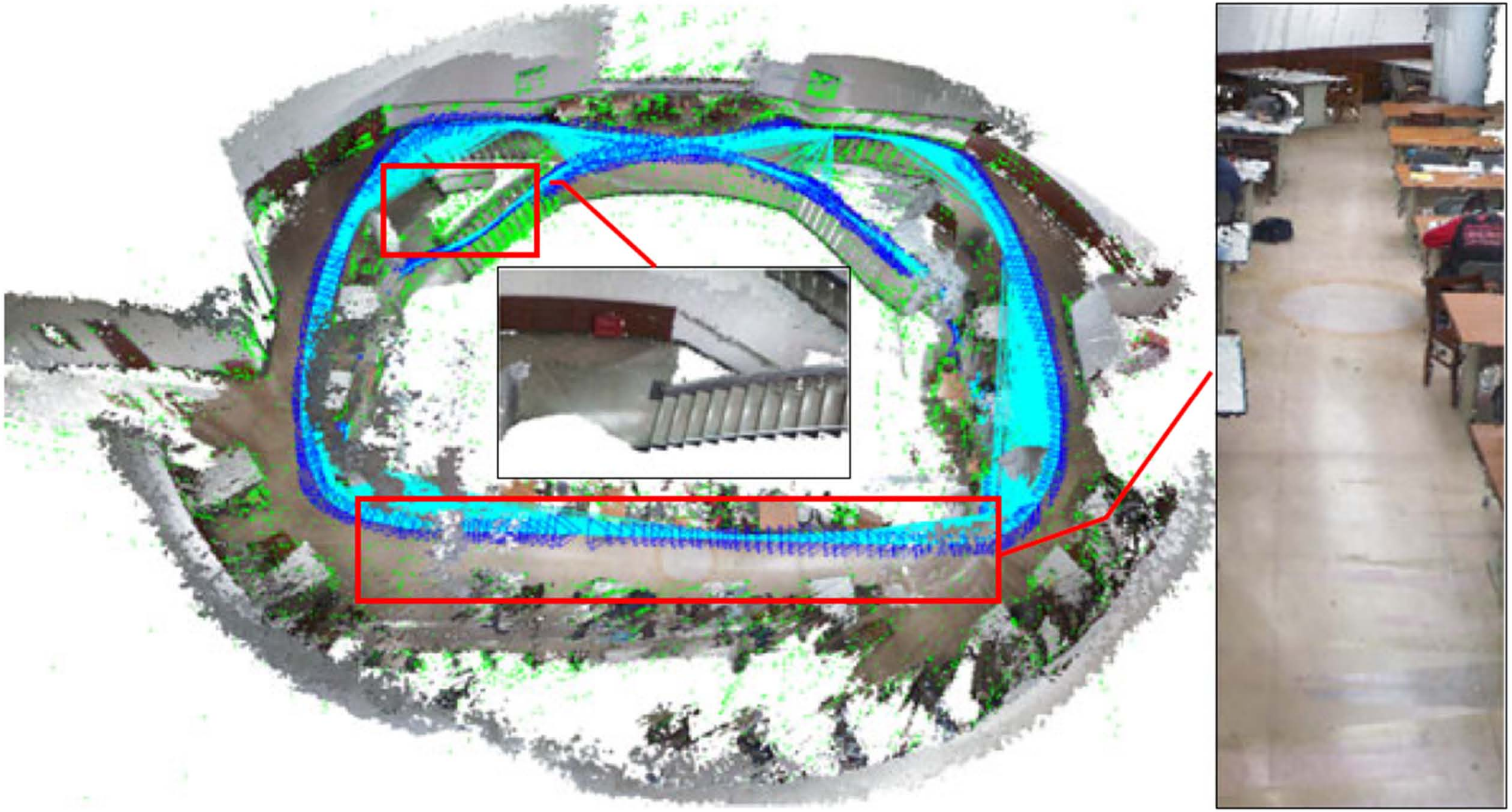
(c) ATE



(d) RPE

Shuhui Bu, Yong Zhao, et al., Semi-direct Tracking and Mapping with RGB-D Camera for MAV, MTAP, 2016.

Experiment Results (NPU Dataset)



Shuhui Bu, Yong Zhao, et al., Semi-direct Tracking and Mapping with RGB-D Camera for MAV, MTAP, 2016.

Comparison

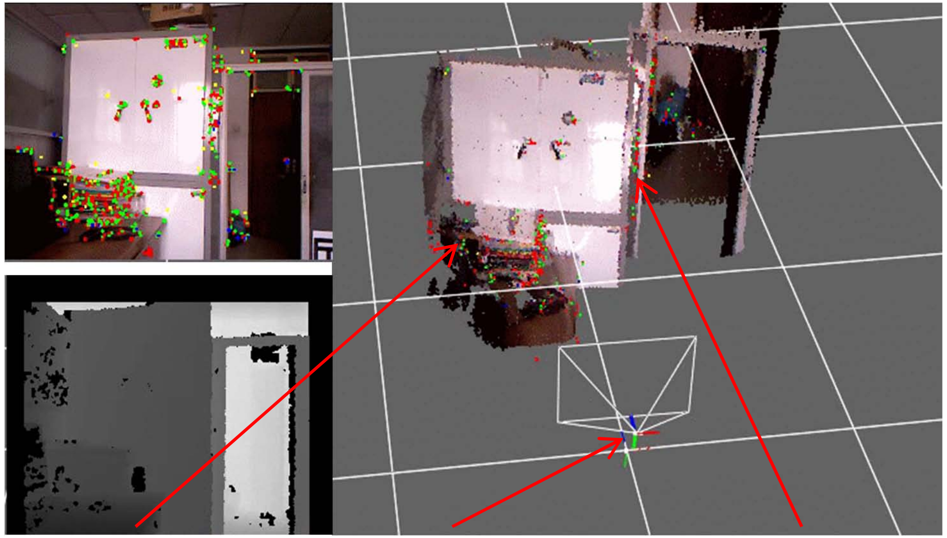
Experiments on TUM dataset

Sequence Name	SDTAM (Ours)			DVO [34]	Kinect Fusion [29]	RGB-D SLAM [25]	Volume Fusion [36]
	Direct	Direct+KF	Direct+KF+Loop				
fr1/xyz	0.054	**0.011**	**0.011**	**0.011**	0.026	0.014	0.017
fr1/rpy	0.086	0.031	0.031	**0.020**	0.133	0.026	0.028
fr1/desk	0.055	**0.018**	**0.018**	0.021	0.057	0.023	0.037
fr1/desk2	0.117	**0.043**	**0.043**	0.046	0.420	0.043	0.071
fr1/room	0.305	0.205	0.084	**0.053**	0.313	0.084	0.075
fr1/plant	0.039	0.072	0.034	**0.028**	0.598	0.091	0.047
fr2/xyz	0.017	0.015	0.015	0.018	-	**0.008**	0.029
fr2/person	0.180	**0.079**	**0.079**	-	-	-	-
fr3/long	0.104	0.018	**0.010**	0.035	0.064	0.032	0.030
fr3/nst	0.045	0.020	**0.013**	0.018	-	0.017	0.031
fr3/far	0.010	**0.009**	**0.009**	0.017	-	-	-
fr3/sit_xyz	0.028	**0.008**	**0.008**	-	-	-	-
fr3/sit_halfsph	0.116	**0.012**	**0.012**	-	-	-	-
fr3/walk_xyz	1.436	**0.011**	**0.011**	-	-	-	-
fr3/walk_halfsph	0.649	**0.060**	**0.060**	-	-	-	-

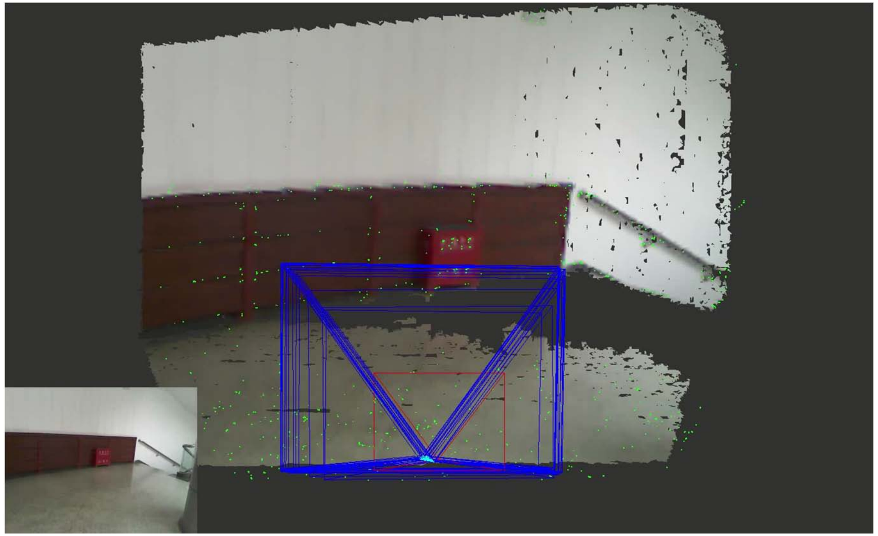
Shuhui Bu, Yong Zhao, et al., Semi-direct Tracking and Mapping with RGB-D Camera for MAV, MTAP, 2016.

Applications

- Realtime, accurate
- GPS-denied navigation
- Indoor robot
- Exploration and rescue



Feature point Keyframe Point cloud



Shuhui Bu, Yong Zhao, et al., Semi-direct Tracking and Mapping with RGB-D Camera for MAV, MTAP, 2016.

NPU RGB-D Dataset

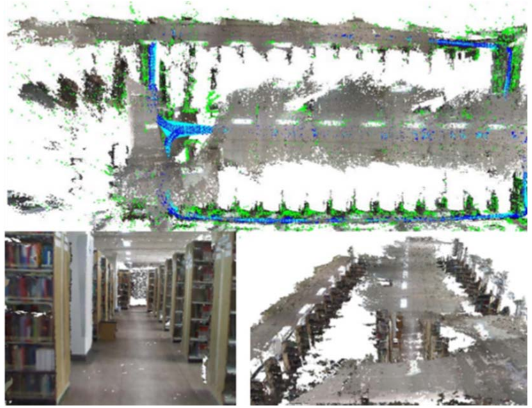
http://www.adv-ci.com/blog/source/npu-rgb-d-dataset/

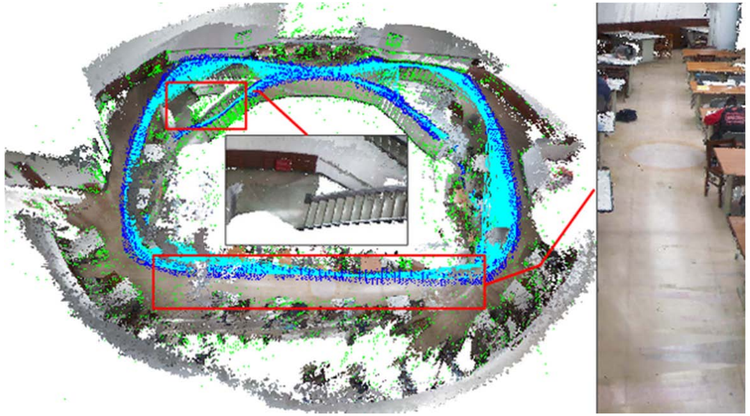
NPU RGB-D Dataset

1. Large-scale RGB-D Dataset

In order to evaluate the RGB-D SLAM algorithms of handling large-scale sequences, we recorded this dataset which contains several sequences in the campus of Northwestern Polytechnical University with a Kinect for XBOX 360. This is a challenging dataset since it contains fast motion, rolling shutter, repetitive scenes, and even poor depth informations.

Each sequence is composed of thousands of colorful and depth image pairs which have been aligned with OpenNI and organized as the format used in TUM dataset. We provide reconstruction results of SDTAM (Semi-direct Tracking and Mapping, a RGB-D localization and mapping algorithm) and sample movies, so that users can preview the sequences before downloading.





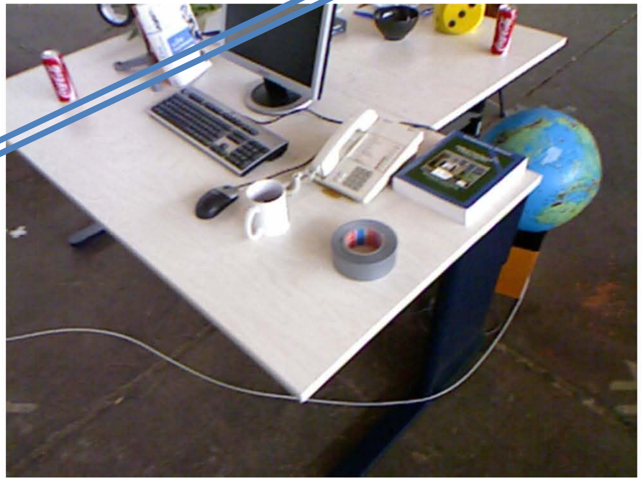
Shuhui Bu, Yong Zhao, et al., Semi-direct Tracking and Mapping with RGB-D Camera for MAV, MTAP, 2016.

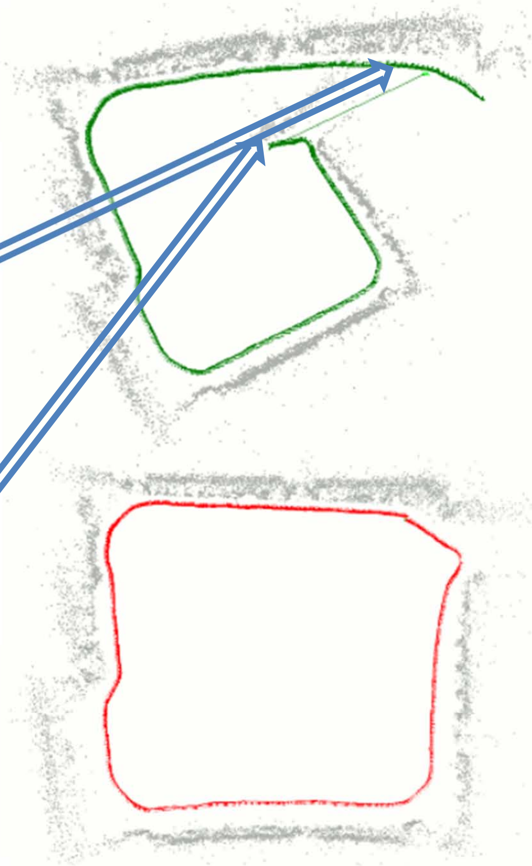
内容

- 研究背景
- 视觉SLAM
  - Map2DFusion： Real-time Mapping for UAV
  - Semi-direct Tracking and Mapping (SDTAM)
  - Place Recognition based on Deep Learning
- 展望未来

Loop Closure

- What is loop closure?
  - To recognize visited places.

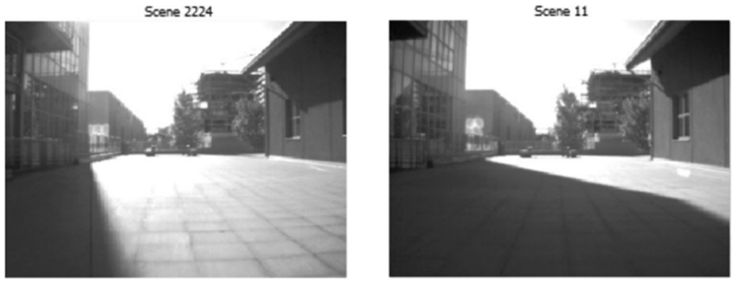




Loop Closure - Challenges

YES

Looks same but not

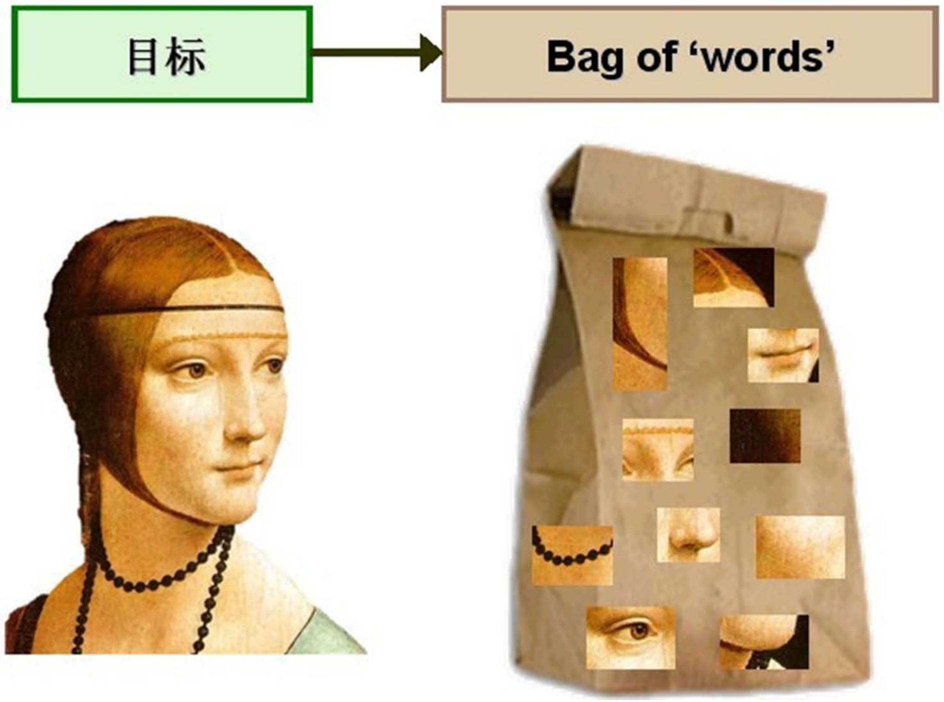


Looks different but the same

False Positive

False Negative

Traditional Way – Bag of Words



From many images: nose, eyes, hairs, …

⇩

Dictionary

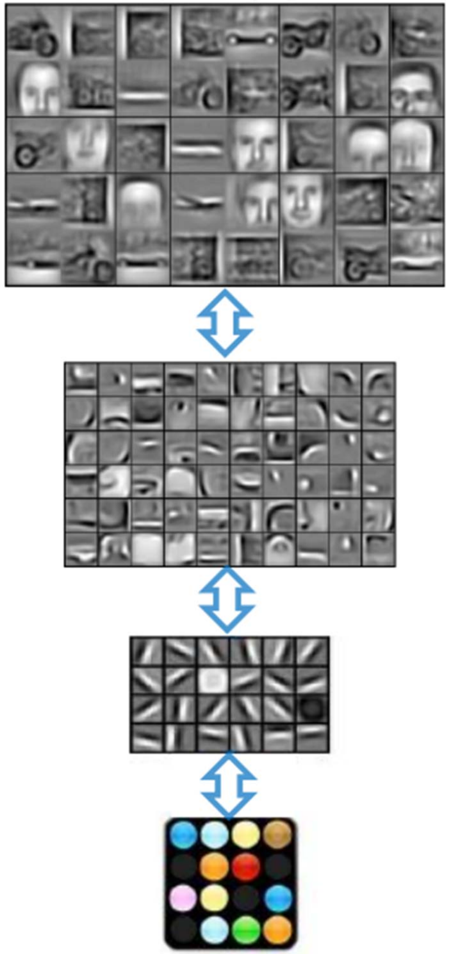
⇩

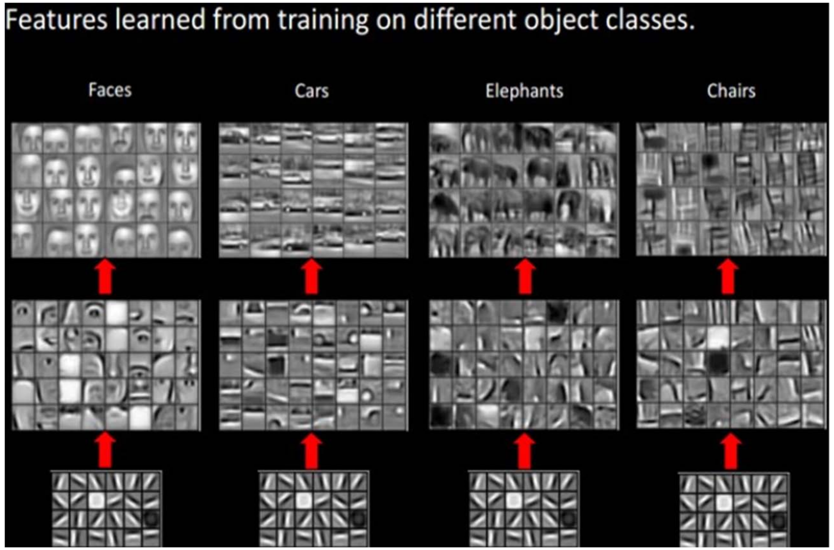
Face = 1 nose + 2 eyes + 1 hairs + …

⇩

Features -> Words

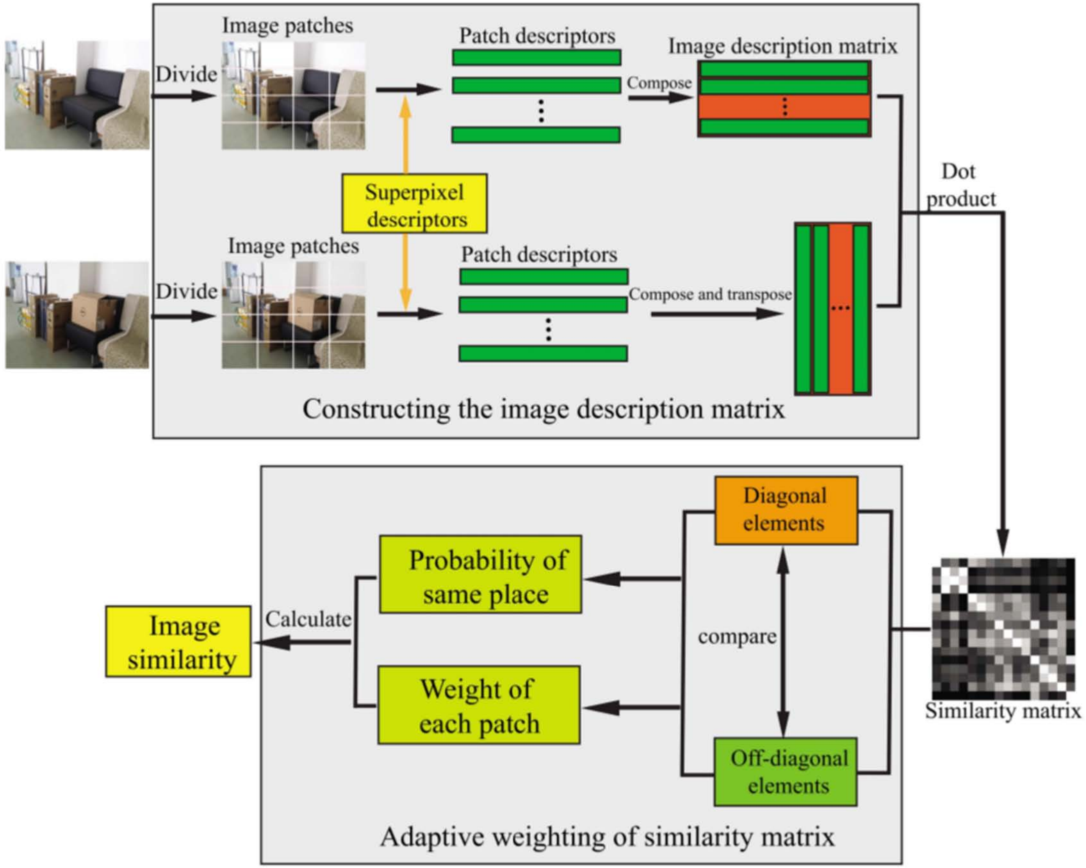
Structural Representation





- Deep architectures can be represented *efficient.*
- *Natural progression* from low level to high level structures.
- Can *share the lower-level* representations for multiple tasks.

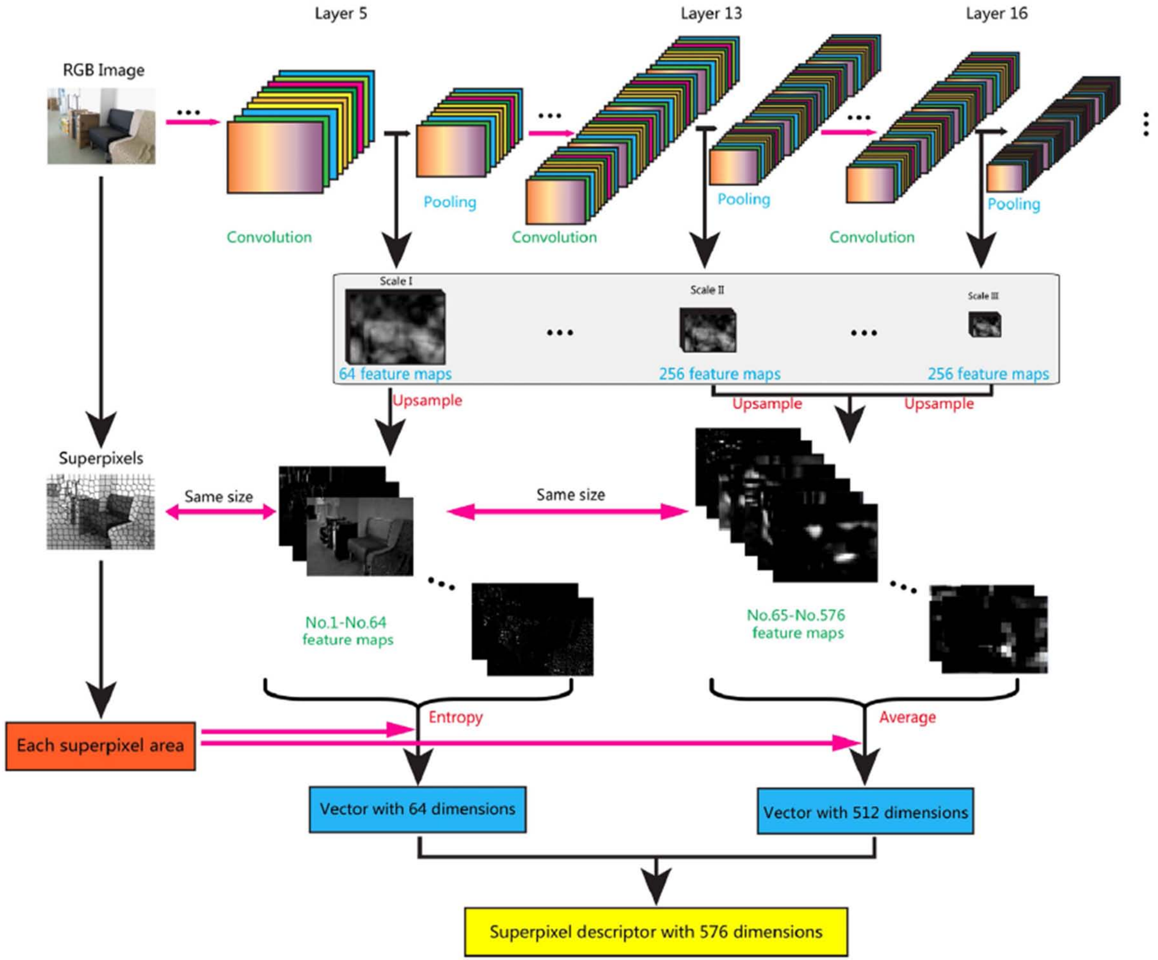
Place Recognition – Deep Way



- Using CNN to extract representative feature
- Considering spatial relationship between objects in scene

Qing Li, et al., Place Recognition Based on Deep Feature and Adaptive Weighting of Similarity Matrix, Neurocomputing, 2016.
Part of source codes: http://www.adv-ci.com/blog/source/pi-cnn and http://www.adv-ci.com/blog/source/pi-slic

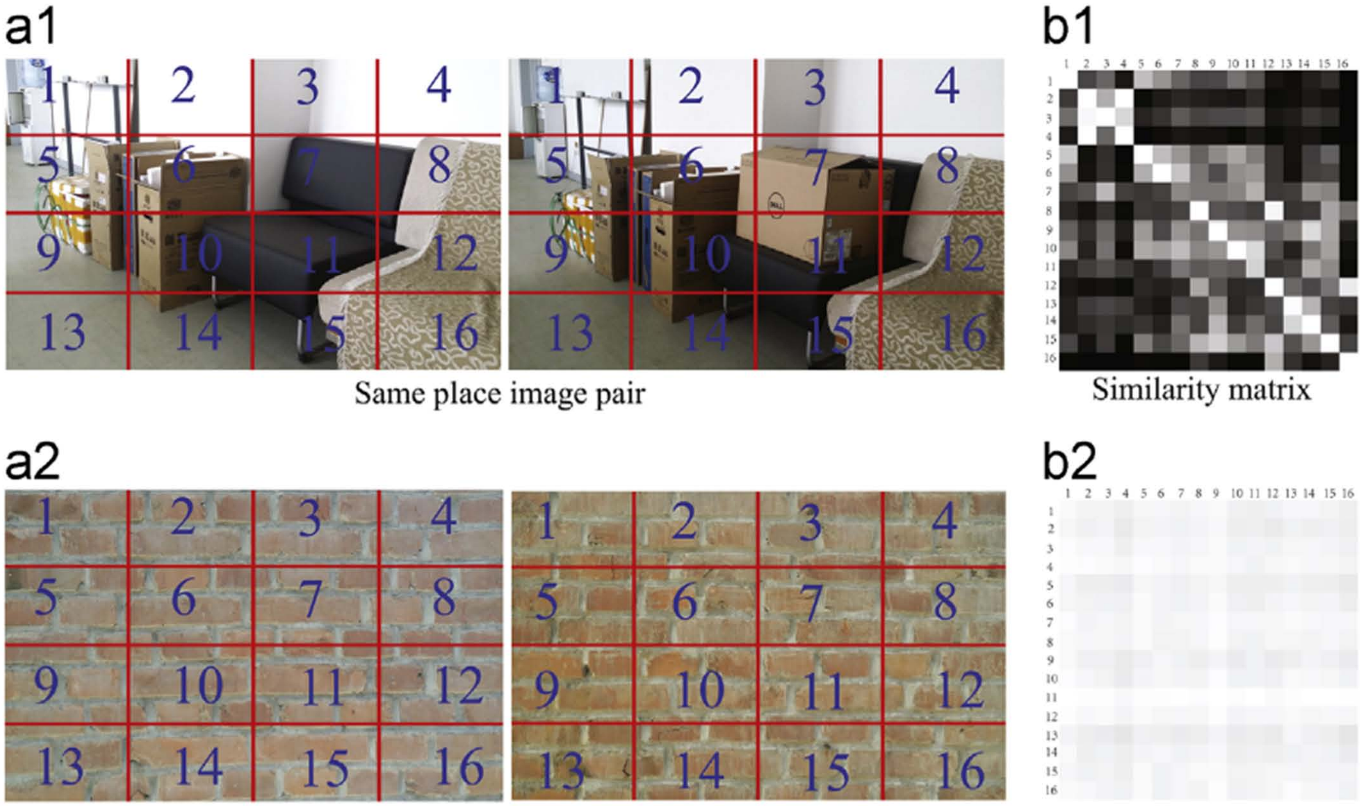
Hierarchical Feature



- The CNN model is 'Imagenet-VGG'
- Select the feature maps of layers 5, 13, and 16.
- Layer 5's superpixel feature is generated by calculating entropy.
- Layer 13 and 16's superpixel features are generated through average.

source codes: http://www.adv-ci.com/blog/source/pi-cnn

Similarity Matrix



Same place image pair

Similarity matrix

Different place image pair

Similarity matrix

$$S_{ij}^{p} = \cos \ < \overrightarrow{F_i^p}, \overrightarrow{F_j^p} > \ = \frac{\overrightarrow{F_i^p} \cdot \overrightarrow{F_j^p}}{|\overrightarrow{F_i^p}||\overrightarrow{F_j^p}|},$$

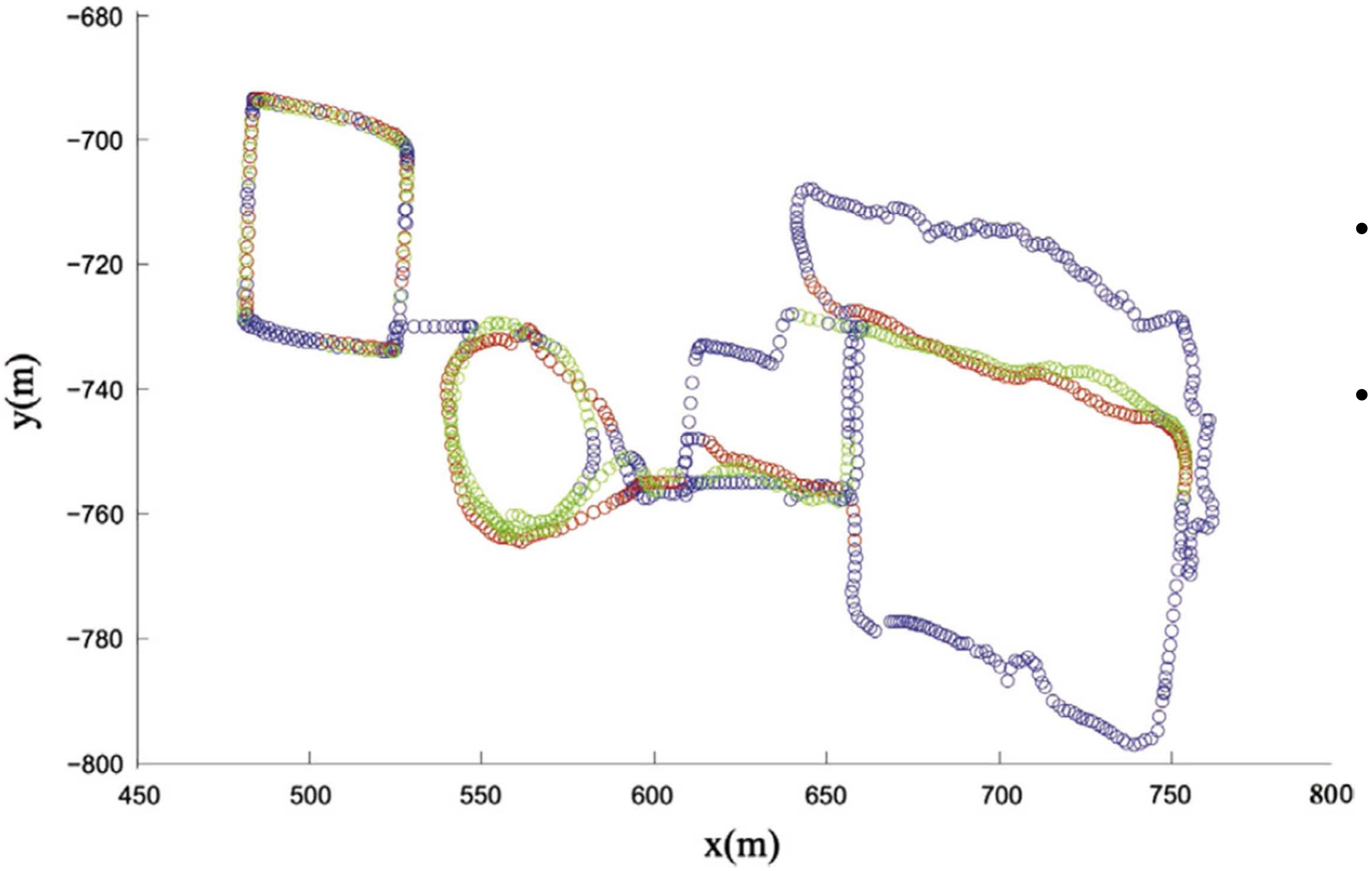
$$S = \xi \sum_{i=1}^{16} w_i^p S_{ii}^p,$$

$$x_i = \begin{cases} \dfrac{d_i - \mu_r}{\delta} & : \ d_i > \mu_r, \\ 0 & : \ d_i \leq \mu_r, \end{cases}$$

$$w_i^p = \frac{x_i}{\sum_{j=1}^{16} x_j}$$

Qing Li, et al., Place Recognition Based on Deep Feature and Adaptive Weighting of Similarity Matrix, Neurocomputing, 2016.

Experiment Results



- The revisited location are labeled by the red circles.
- Similar with revised ones are labeled with green circle.

Qing Li, et al., Place Recognition Based on Deep Feature and Adaptive Weighting of Similarity Matrix, Neurocomputing, 2016.

内容

- 研究背景
- 视觉SLAM
  - Map2DFusion： Real-time Mapping for UAV
  - Semi-direct Tracking and Mapping (SDTAM)
  - Place Recognition based on Deep Learning
- 展望未来

未来之路

人工智能

环境感知

推理

决策

SLAM

场景理解

多数据源
图像
LiDAR
红外
多光谱，高光谱
IMU
MAP

实时，低延时

复杂环境
电磁环境，气象环境

- 鲁棒性高、精度高的SLAM
- 语义层面的地图
- 基于深度学习的场景理解方法
- 推理、策略的学习
- 智能多机协同系统

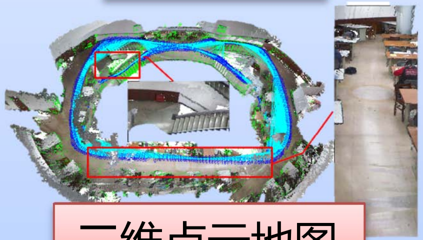
地图与机器

- 地图不仅仅是为人提供空间信息的工具
- 更多是为机器提供空间信息，导航定位、环境感知
- 智能机器不仅仅是地图的使用者，也是地图的生成者，两者相互依赖
- 要求更高的实时性，更多样的表达形式

智能机器

机器地图

影像地图

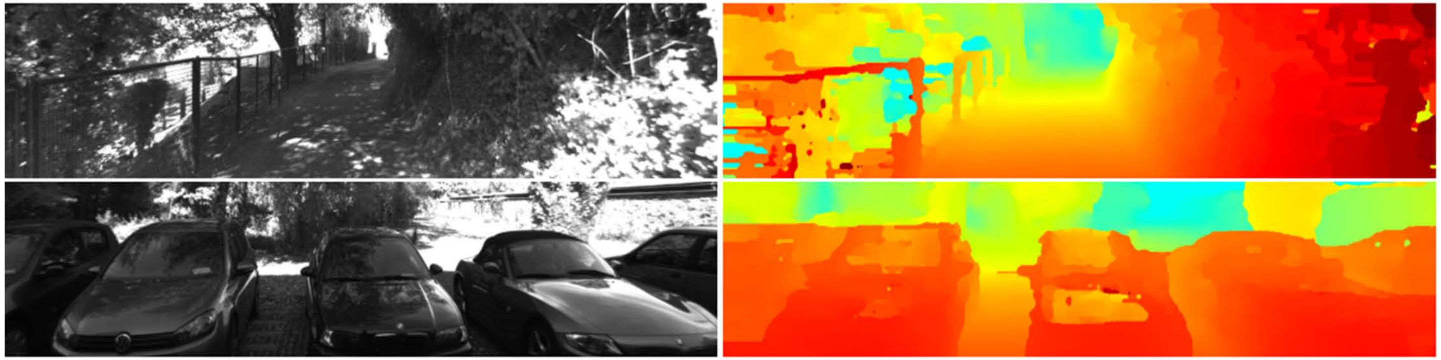


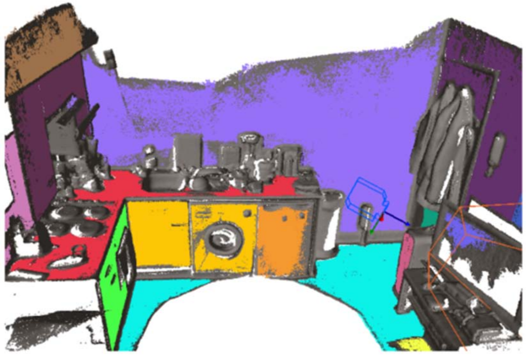
三维点云地图

语义地图

SLAM的发展方向

- 语义层面的信息能够实现更大规模、更可靠的SLAM系统
- 基于直接优化的方法，基于深度学习的光流
- SLAM为深度学习提供学习数据
- 端到端的学习实现导航、制导，全神经网络的机器人系统





THANK YOU

Detailed information can be found at http://www.adv-ci.com