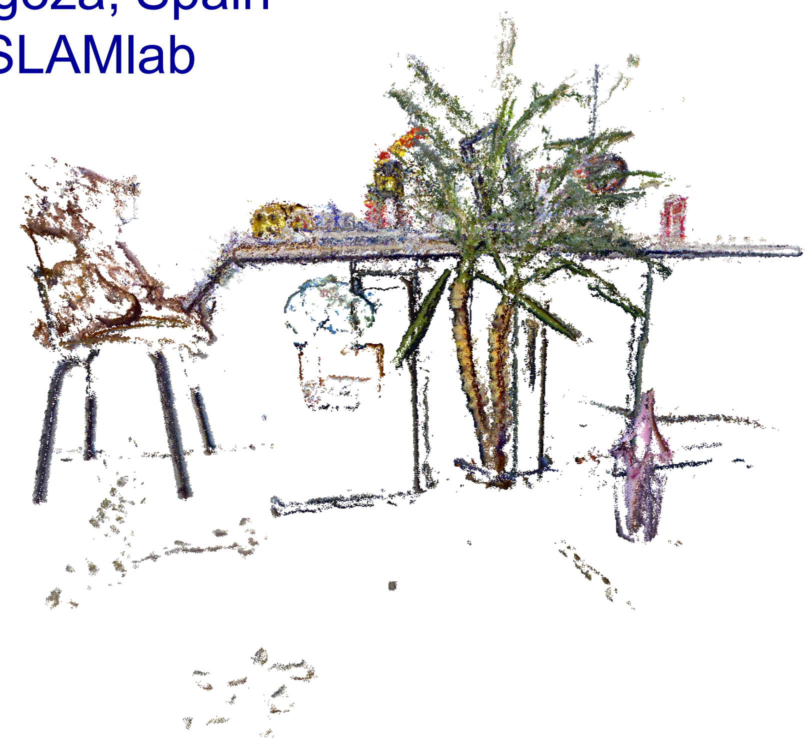
Qualcomm Augmented Reality Lecture Series
Vienna - June 11, 2015

ORB-SLAM: a Real-Time Accurate Monocular SLAM System

Juan D. Tardós, Raúl Mur Artal, José M.M. Montiel
Universidad de Zaragoza, Spain
robots.unizar.es/SLAMlab



Outline

- Motivation
- ORB-SLAM: Feature-Based Visual SLAM
- Robustness and Accuracy
- Application to VR
- Discussion and Future Work

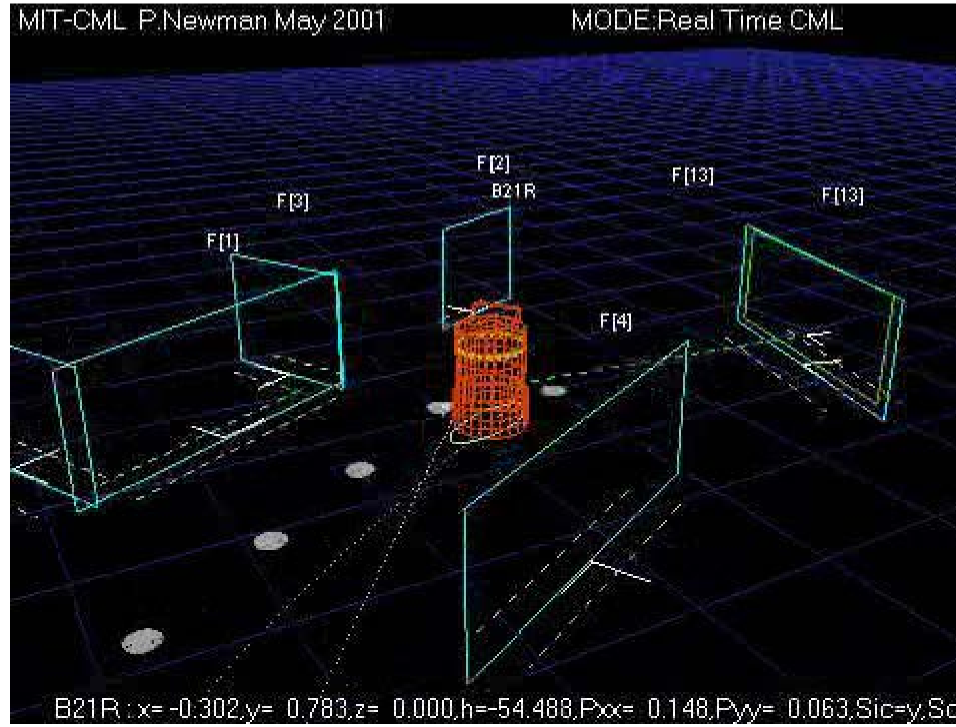
SLAM: Simultaneous Localization and Mapping

The SLAM problem:

- a robot moving in an unknown environment

Use sensor data to:

- **build a map** of the environment
- and **at the same time**
- use the map to compute the **robot location**

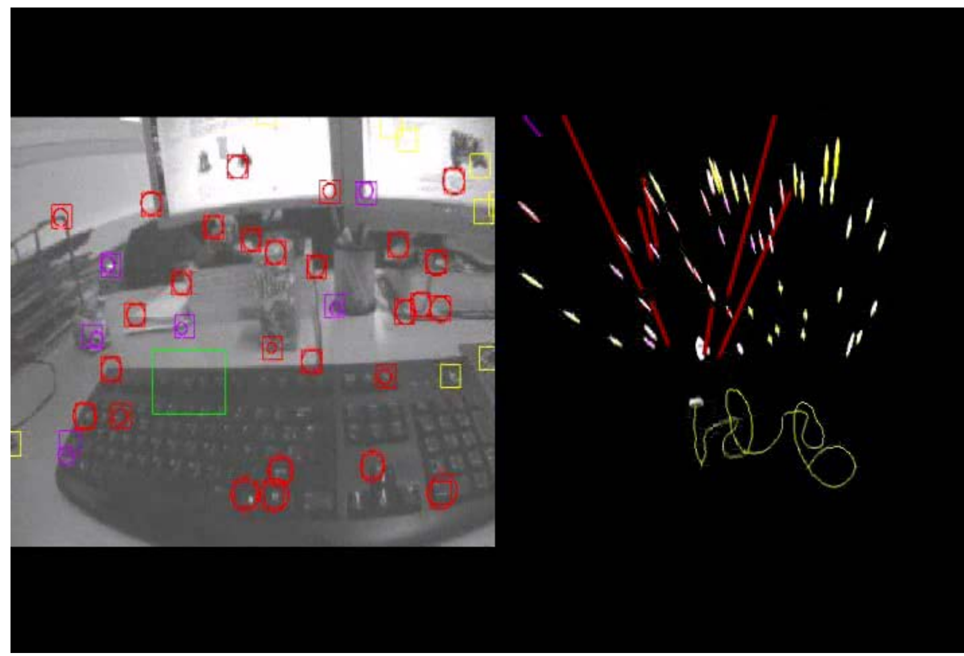


P. Newman, J.J Leonard, J.D. Tardos, J. Neira: **Explore and return: Experimental validation of real-time concurrent mapping and localization.** IEEE Int. Conf. Robotics and Automation, 2002

Motivation

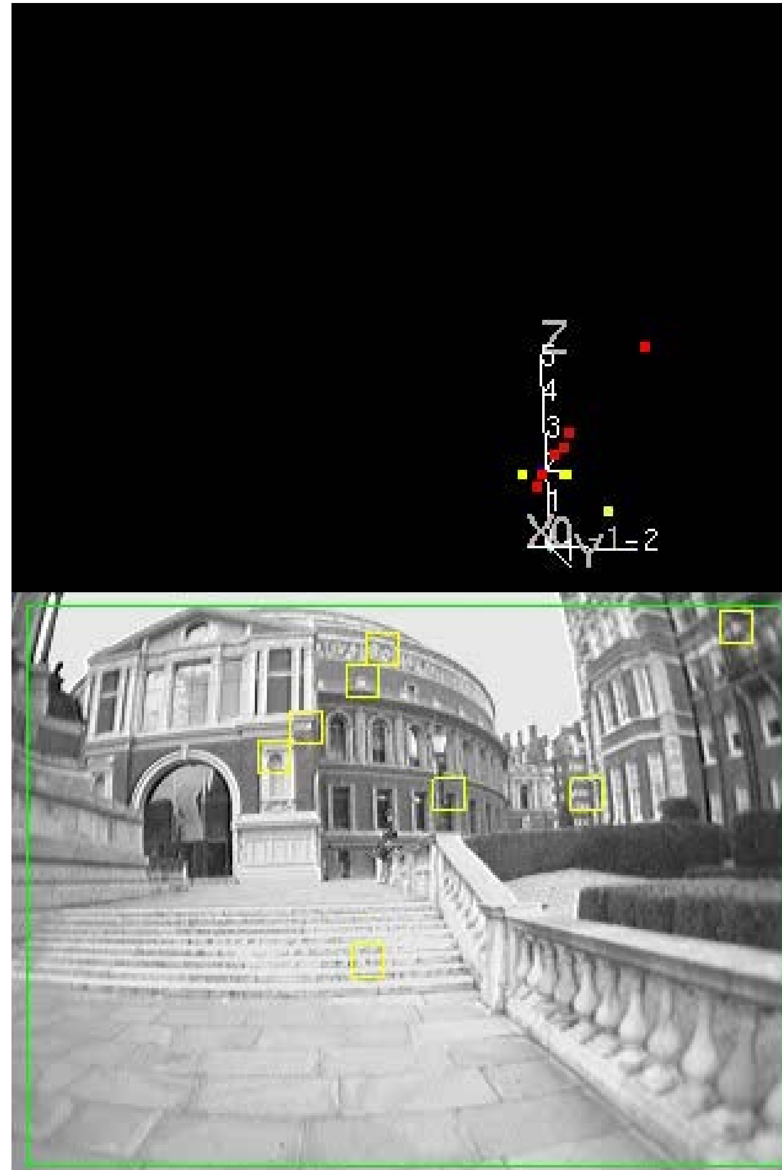
- Head tracking in VR
  - Orientation tracking
  - Based on IMU
  - Slow drift
  - Position tracking with external camera
    - » Limited by the field of view

- Real-time Visual SLAM
  - 6DOF camera tracking
  - Monocular, stereo, RGB-D
  - Zero drift
  - Unlimited area
  - And gives a Map

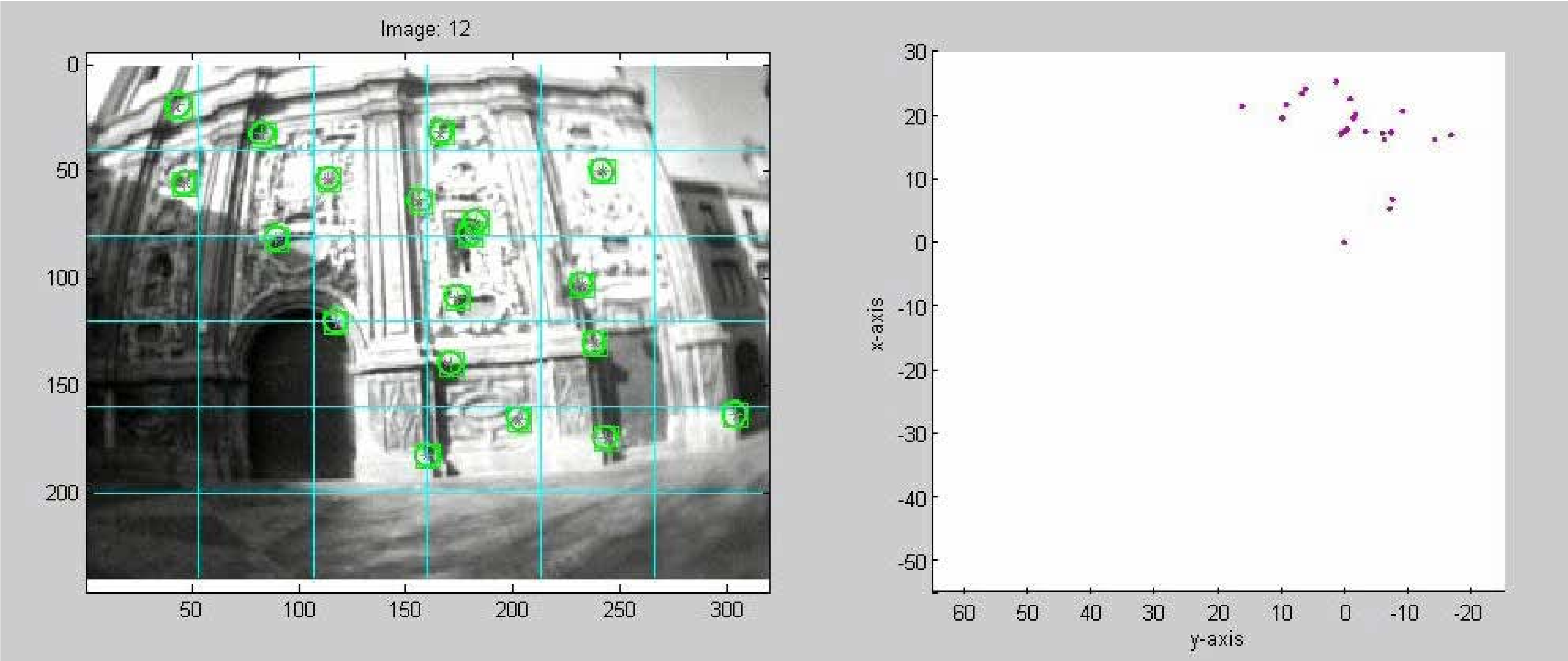


Monocular SLAM with EKF

L.A. Clemente, A.J. Davison, I.D. Reid, J Neira, J.D. Tardós, **Mapping Large Loops with a Single Hand-Held Camera.** Robotics: Science and Systems, June 2007



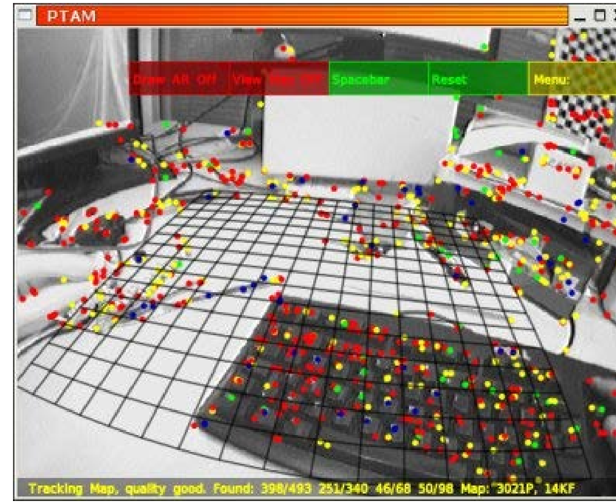
Monocular SLAM with EKF



P. Piniés, J.D. Tardós, **Large-scale SLAM building conditionally independent local maps: Application to monocular vision**, IEEE Transactions on Robotics, 24 (5), 1094-1106, 2008

PTAM: Keyframe-Based SLAM

- *Parallel Tracking and Mapping for Small AR Workspaces G. Klein and D. Murray, ISMAR 2007*
  - First Keyframe-Based Monocular SLAM
  - Bundle Adjustment is possible in real-time
  - Still considered a reference system

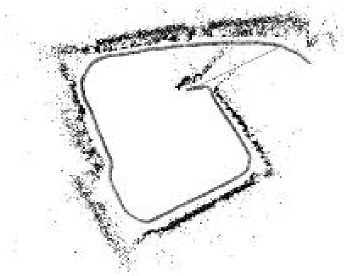


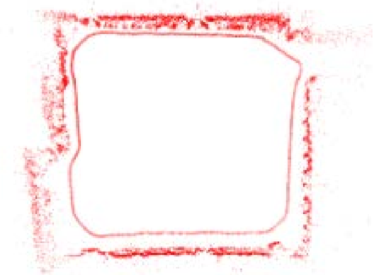
- Limitations:
  - Relocalisation with little invariance to viewpoint
  - No loop detection
  - Small scale operation
  - Initialization in planar scenes, with human intervention

Recent Key Ideas

- Scale Drift-Aware Loop Closing

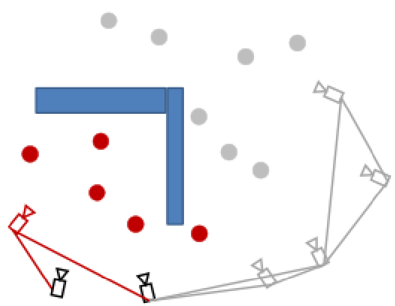
  H. Strasdat, J.M.M. Montiel and A.J. Davison
  Scale Drift-Aware Large Scale Monocular SLAM
  RSS 2010





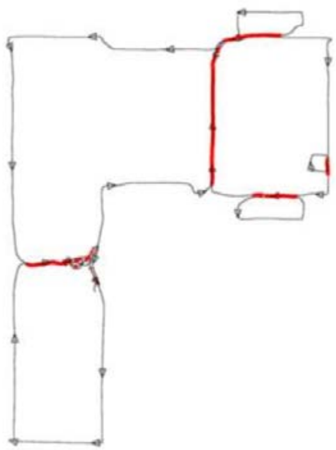
- Covisibility Graph

  H. Strasdat, A. J. Davison, J. M. M. Montiel , K. Konolige
  Double Window Optimization for Constant Time Visual SLAM
  ICCV 2011

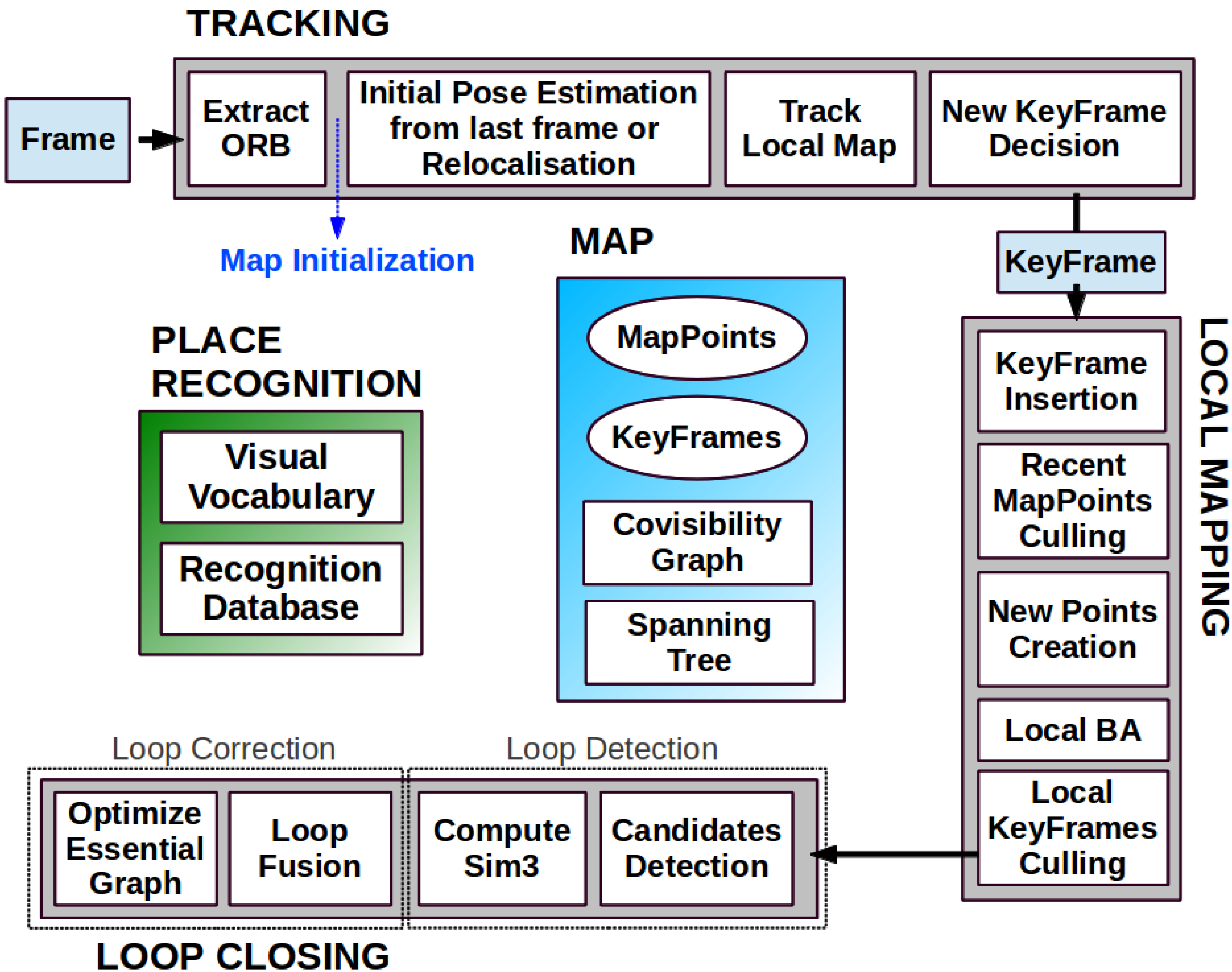


- Bags of Binary Words (DBoW)

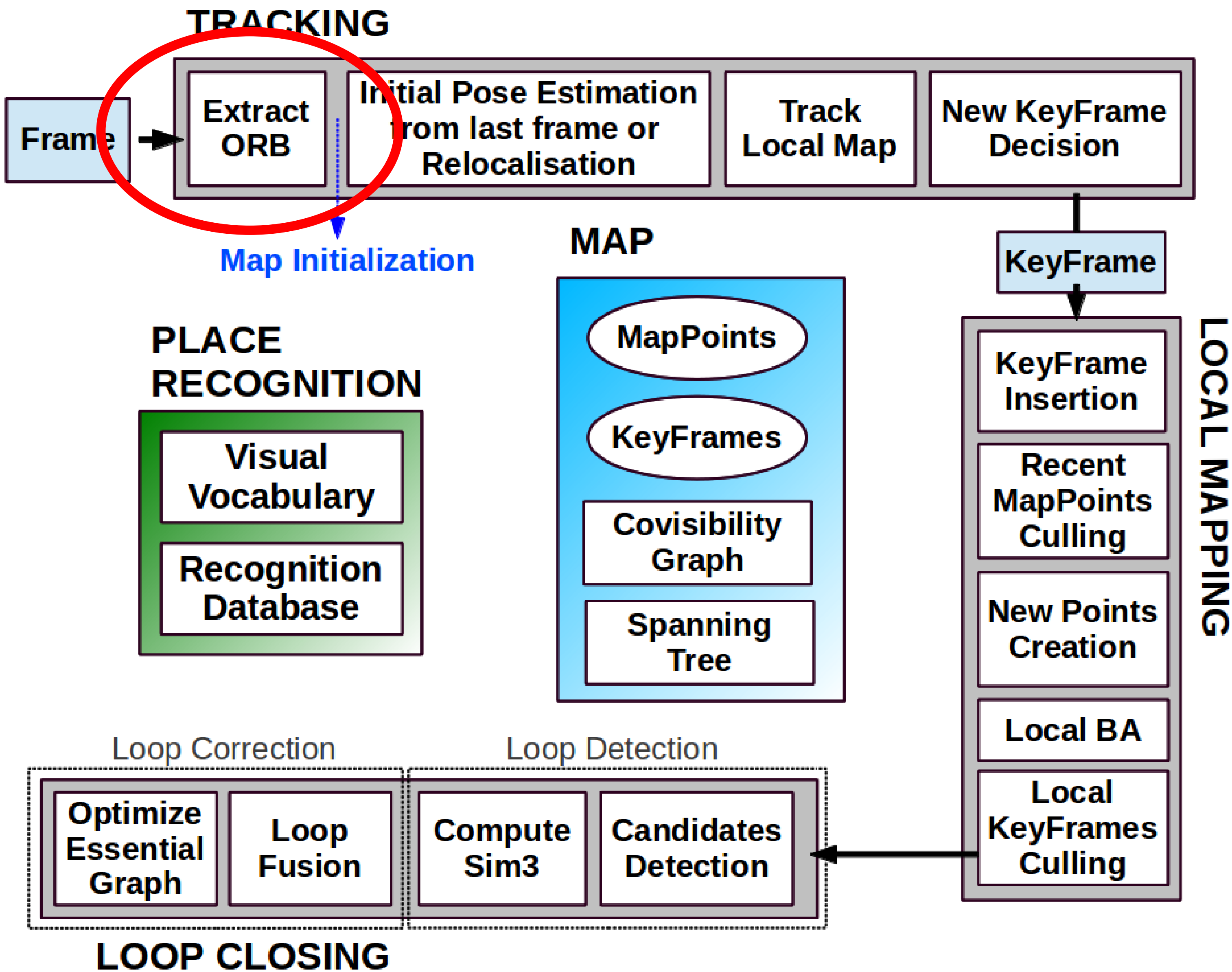
  D. Gálvez-López and J. D. Tardós
  Bags of Binary Words for Fast Place Recognition in Image Sequences, IEEE Transactions on Robotics 2012



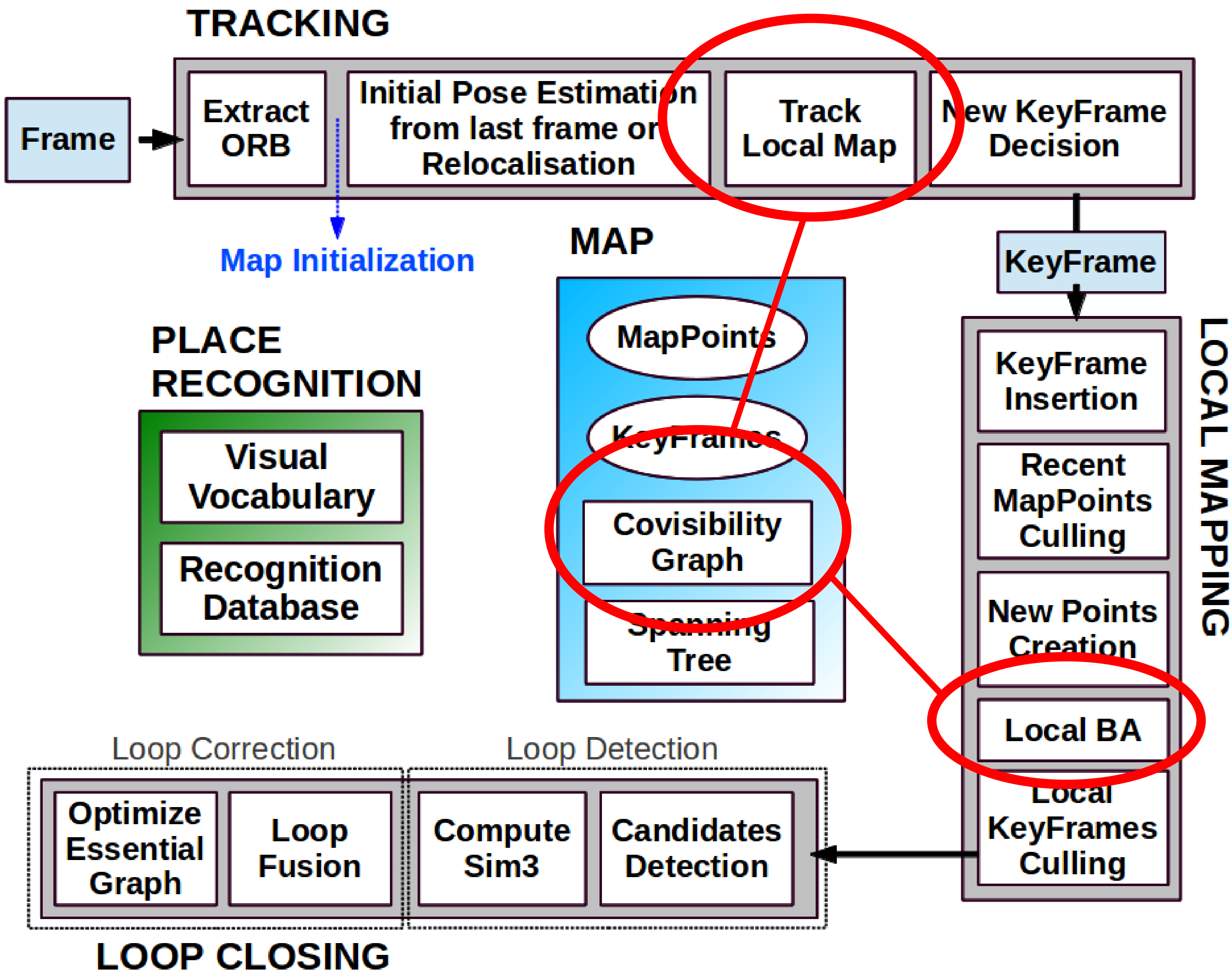
ORB-SLAM: Real-Time Monocular SLAM



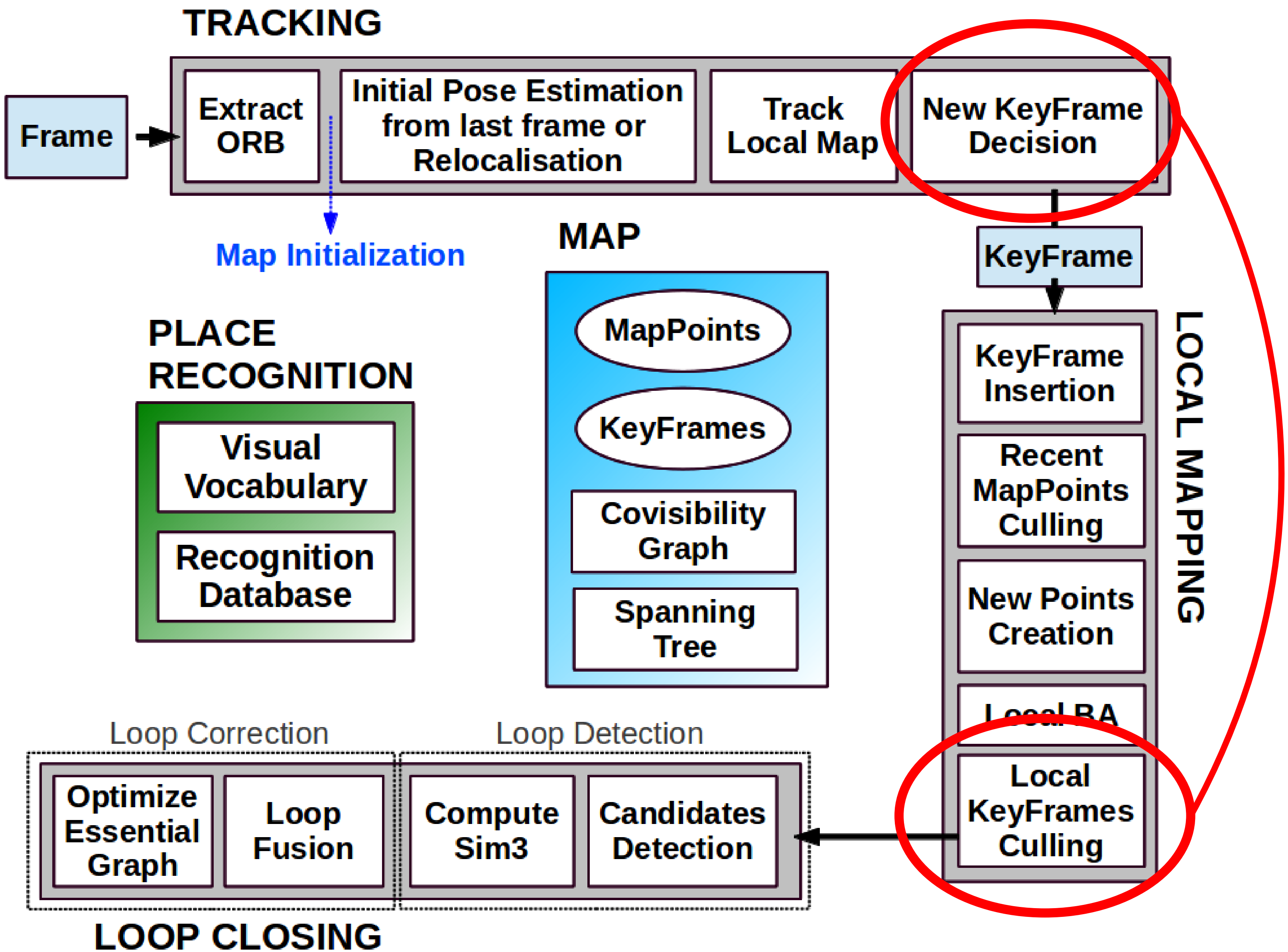
ORB-SLAM: Real-Time Monocular SLAM



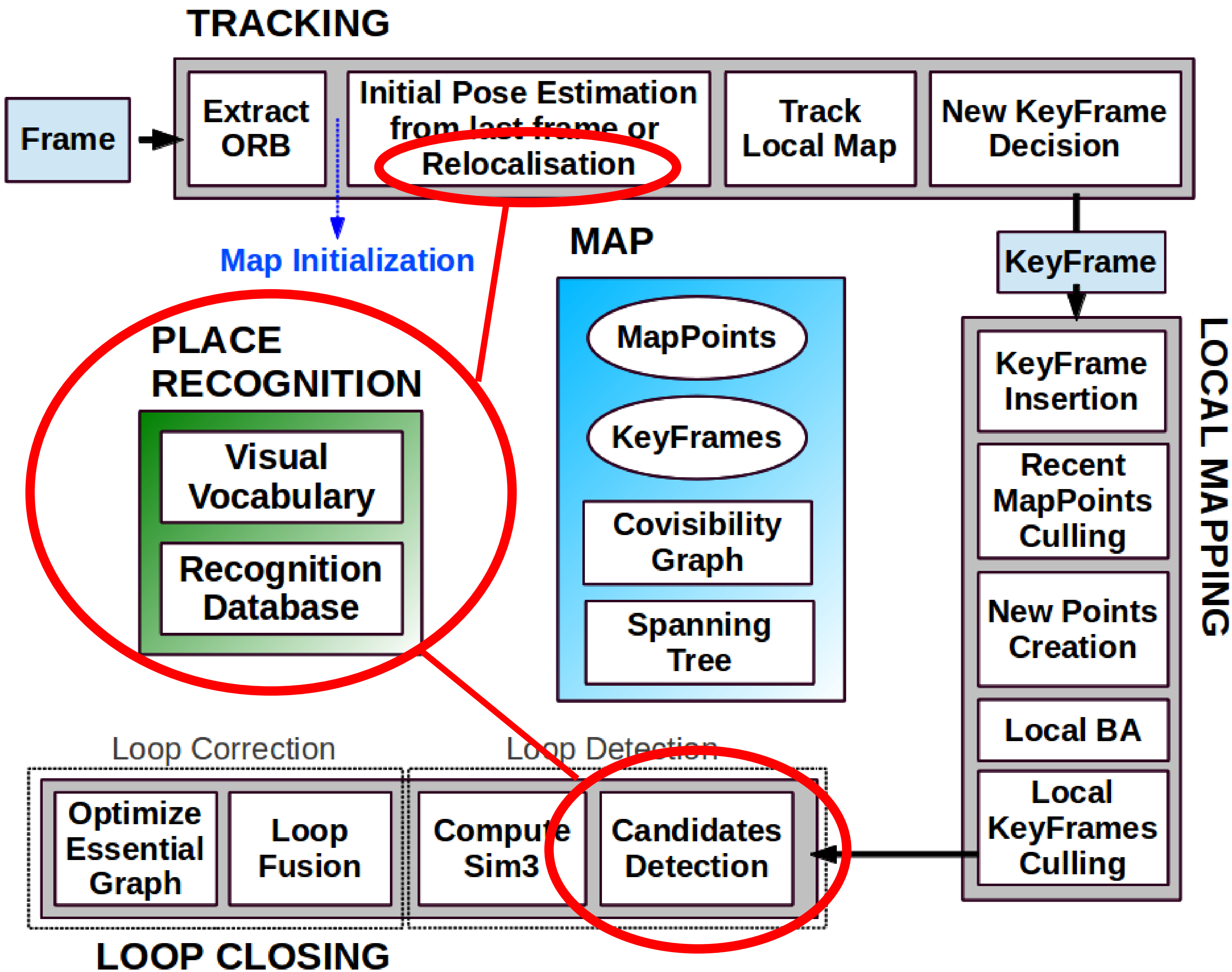
ORB-SLAM: Real-Time Monocular SLAM



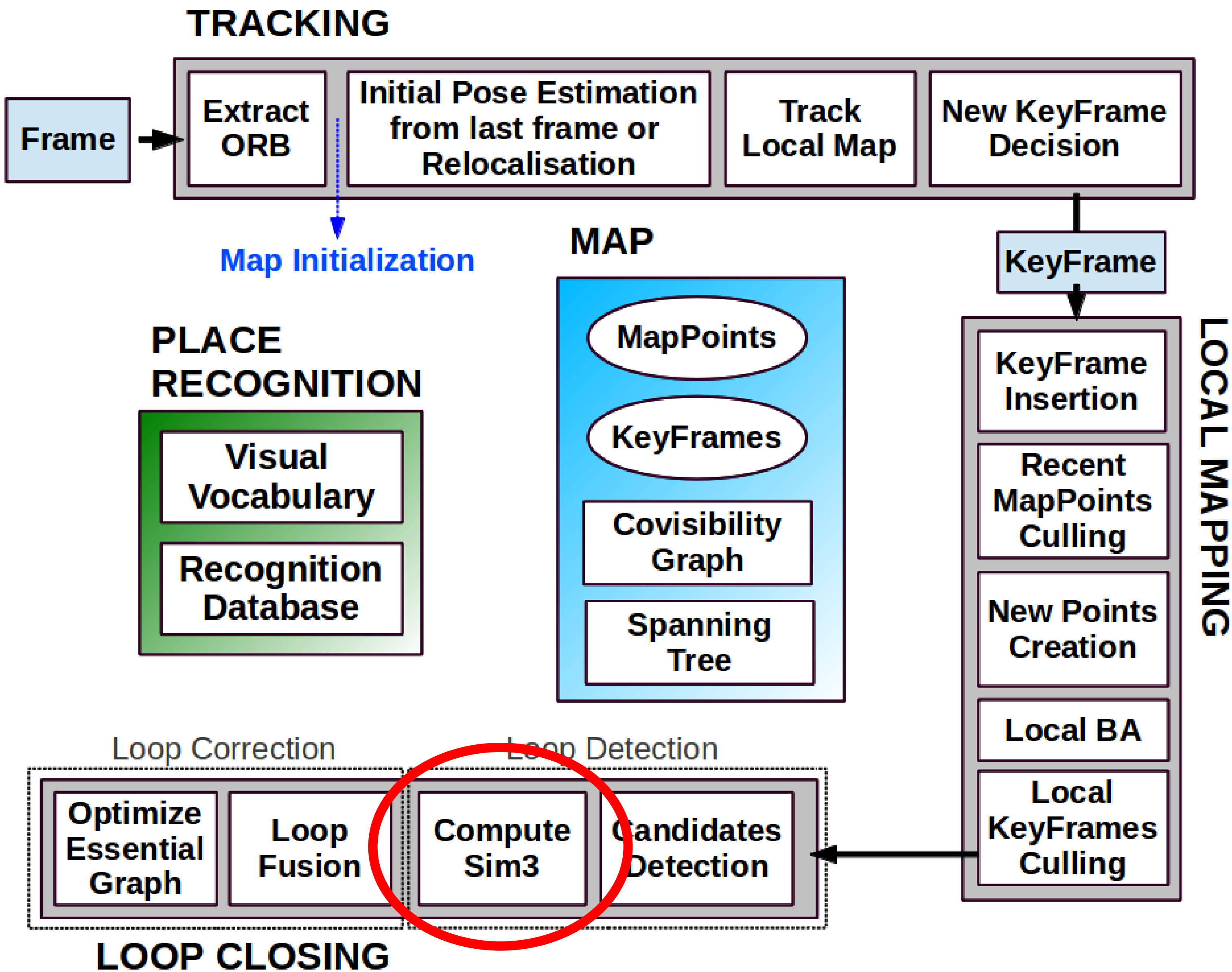
ORB-SLAM: Real-Time Monocular SLAM



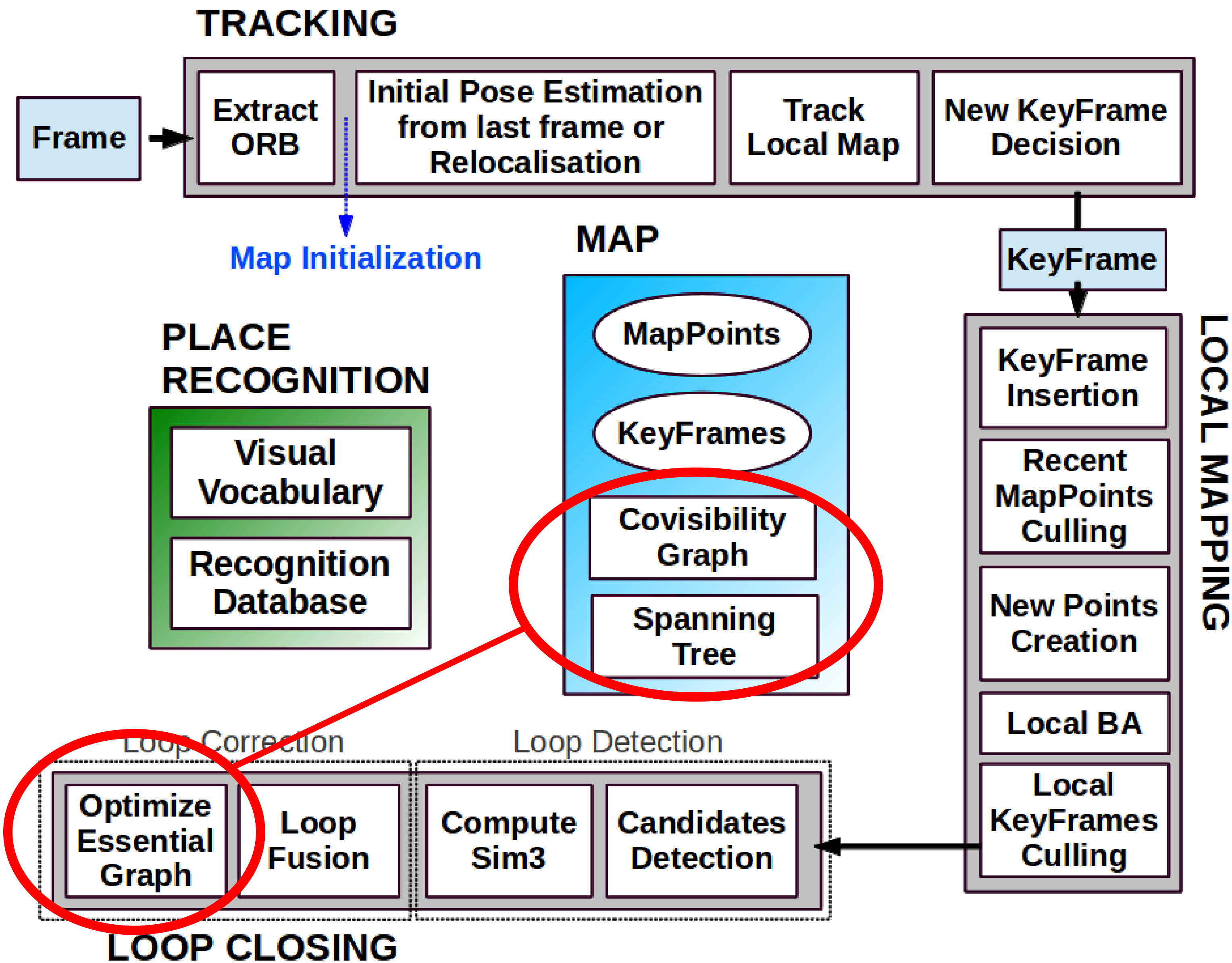
ORB-SLAM: Real-Time Monocular SLAM



ORB-SLAM: Real-Time Monocular SLAM



ORB-SLAM: Real-Time Monocular SLAM



Automatic Map Initialization

Model Selection

Homography
(Planar, Low Parallax)

Fundamental Matrix
(General)

ORB-SLAM

PTAM

Tracking I: Occlusion Handling

<u>Covisibility Graph</u>

Track only a local map (potentially visible)

<u>Point Viewing Direction</u>

Do not project if further than 60º

ORB-SLAM

PTAM

Tracking II: Fast KeyFrame Insertion

Survival of the Fittest
KeyFrame Selection

Fast Keyframe Insertion
(no distance threshold)

Culling of redundant
Keyframes

ORB-SLAM

PTAM

Relocalisation

Bags of Binary Words

Same ORBs used in Tracking and Mapping

Good Viewpoint Invariance (ORB)

ORB-SLAM

PTAM

ORB-SLAM indoors: TUM RGB-D dataset

	ORB-SLAM	PTAM
fr1_xyz	**0.90**	1.15
fr2_xyz	0.30	**0.20**
fr1_floor	**2.99**	
fr1_desk	**1.69**	
fr2_360_kidnap	3.81	**2.63**
fr2_desk	**0.88**	
fr3_long_office	**3.45**	
fr3_nstr_tex_near	**1.39**	2.74
fr3_str_tex_far	**0.77**	0.93
fr3_str_tex_near	1.58	**1.04**
fr2_desk_person	**0.63**	
fr3_sit_xyz	**0.79**	0.83
fr3_sit_halfsph	**1.34**	
fr3_walk_xyz	**1.24**	
fr3_walk_halfsph	**1.74**	

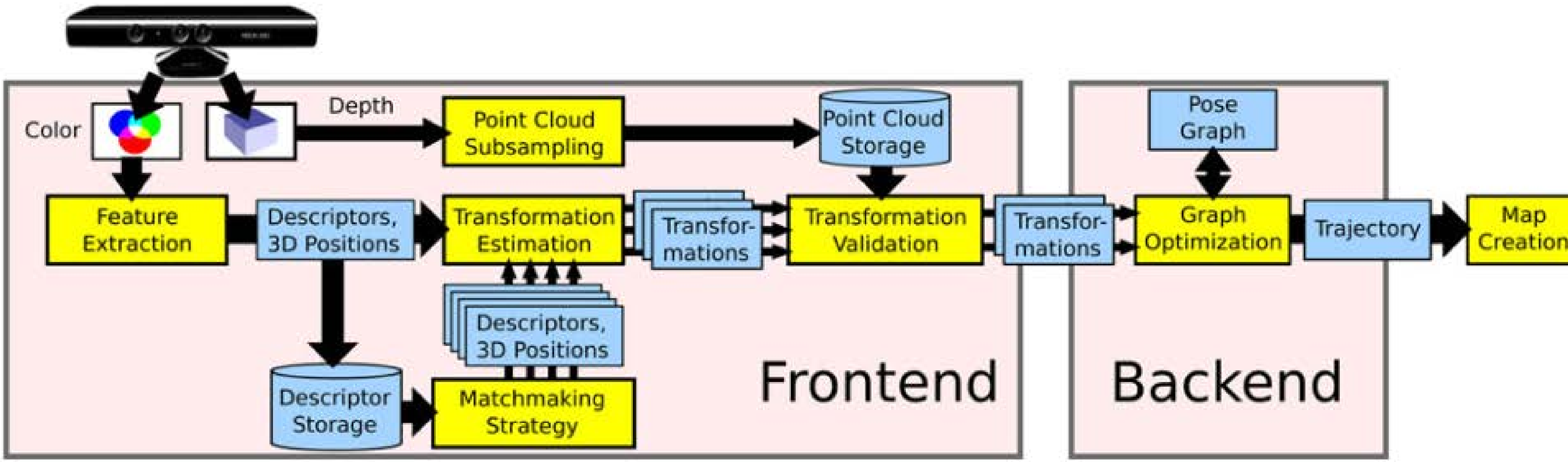
TUM RGB-D Benchmark

RMS KeyFrame Position Error (cm)

Median over 5 executions

Comparison with RGB-D SLAM

Use depth information!



F. Endres, J. Hess, J. Sturm, D. Cremers and W. Burgard
3-D Mapping with an RGB-D Camera
IEEE Transaction on Robotics, 2014.

	ORB-SLAM	PTAM	RGBD-SLAM
fr1_xyz	**0.90**	1.15	1.34
fr2_xyz	0.30	**0.20**	1.42
fr1_floor	**2.99**		3.51
fr1_desk	**1.69**		2.52
fr2_360 _kidnap	3.81	**2.63**	100.5
fr2_desk	**0.88**		3.94
fr3_long _office	**3.45**		-
fr3_nstr_ tex_near	**1.39**	2.74	-
fr3_str_ tex_far	**0.77**	0.93	-
fr3_str_ tex_near	1.58	**1.04**	-
fr2_desk _person	**0.63**		2.00
fr3_sit_ xyz	**0.79**	0.83	-
fr3_sit_ _halfsph	**1.34**		-
fr3_walk _xyz	**1.24**		-
fr3_walk _halfsph	**1.74**		-

TUM RGB-D Benchmark

RMSE (cm)

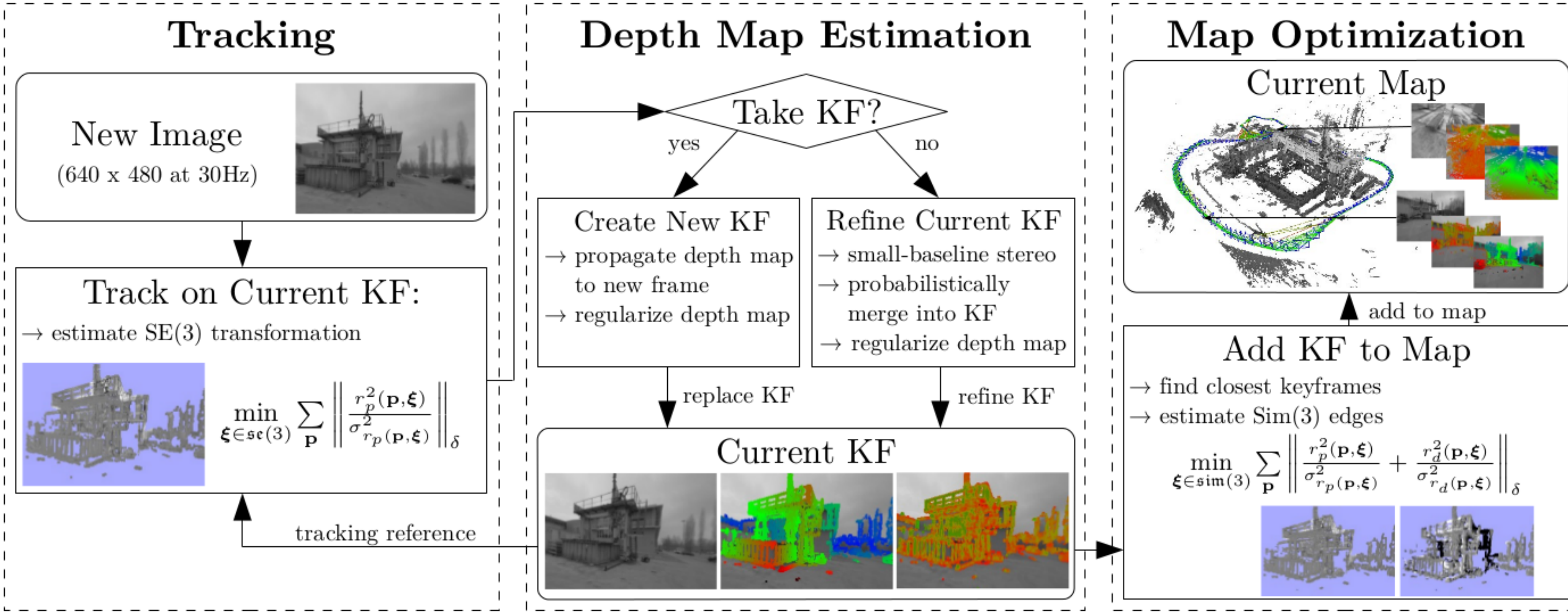
RGB-D SLAM results taken from the benchmark website

ROBOTICS UNIVERSIDAD DE ZARAGOZA

Universidad Zaragoza

Comparison with LSD-SLAM

Use directly pixel intensities!



J. Engel, T. Schöps, D. Cremers
LSD-SLAM: Large-Scale Direct Monocular SLAM
European Conference on Computer Vision (ECCV), 2014.

State-of-the-art in Direct SLAM

Comparison with LSD-SLAM

ORB-SLAM

fr3/structure_texture_near

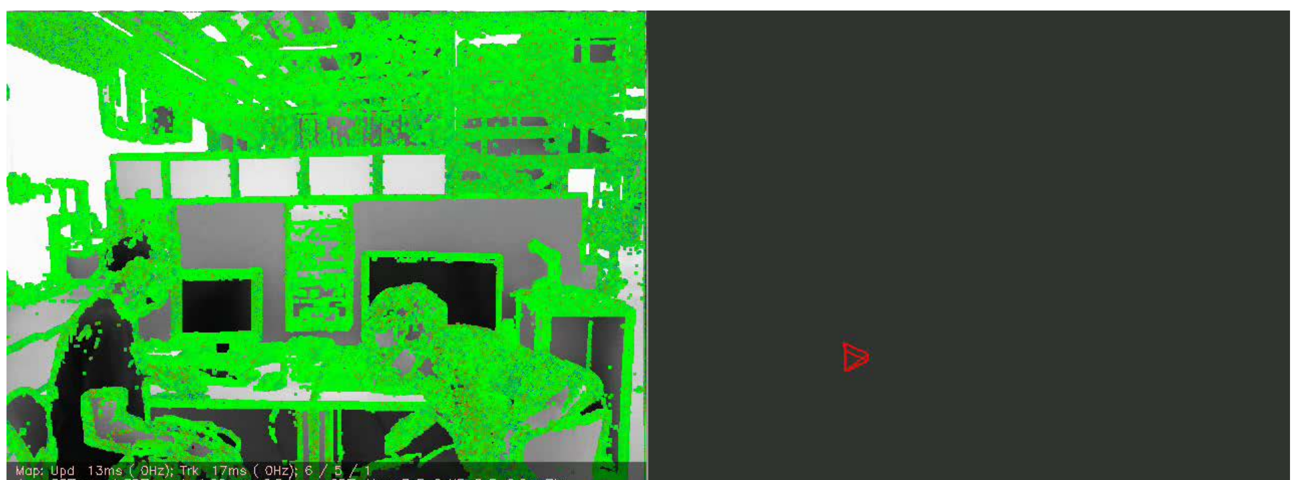
LSD-SLAM

Comparison with LSD-SLAM

ORB-SLAM

fr3/walking_
halfsphere

LSD-SLAM



	ORB-SLAM	PTAM	RGBD-SLAM	LSD-SLAM
fr1_xyz	**0.90**	1.15	1.34	9.00
fr2_xyz	0.30	**0.20**	1.42	2.15
fr1_floor	**2.99**		3.51	38.07
fr1_desk	**1.69**		2.52	10.65
fr2_360_kidnap	3.81	**2.63**	100.5	
fr2_desk	**0.88**		3.94	4.57
fr3_long_office	**3.45**		-	38.53
fr3_nstr_tex_near	**1.39**	2.74	-	7.54
fr3_str_tex_far	**0.77**	0.93	-	7.95
fr3_str_tex_near	1.58	**1.04**	-	
fr2_desk_person	**0.63**		2.00	31.73
fr3_sit_xyz	**0.79**	0.83	-	7.73
fr3_sit_halfsph	**1.34**		-	5.87
fr3_walk_xyz	**1.24**		-	12.44
fr3_walk_halfsph	**1.74**		-	

TUM RGB-D Benchmark

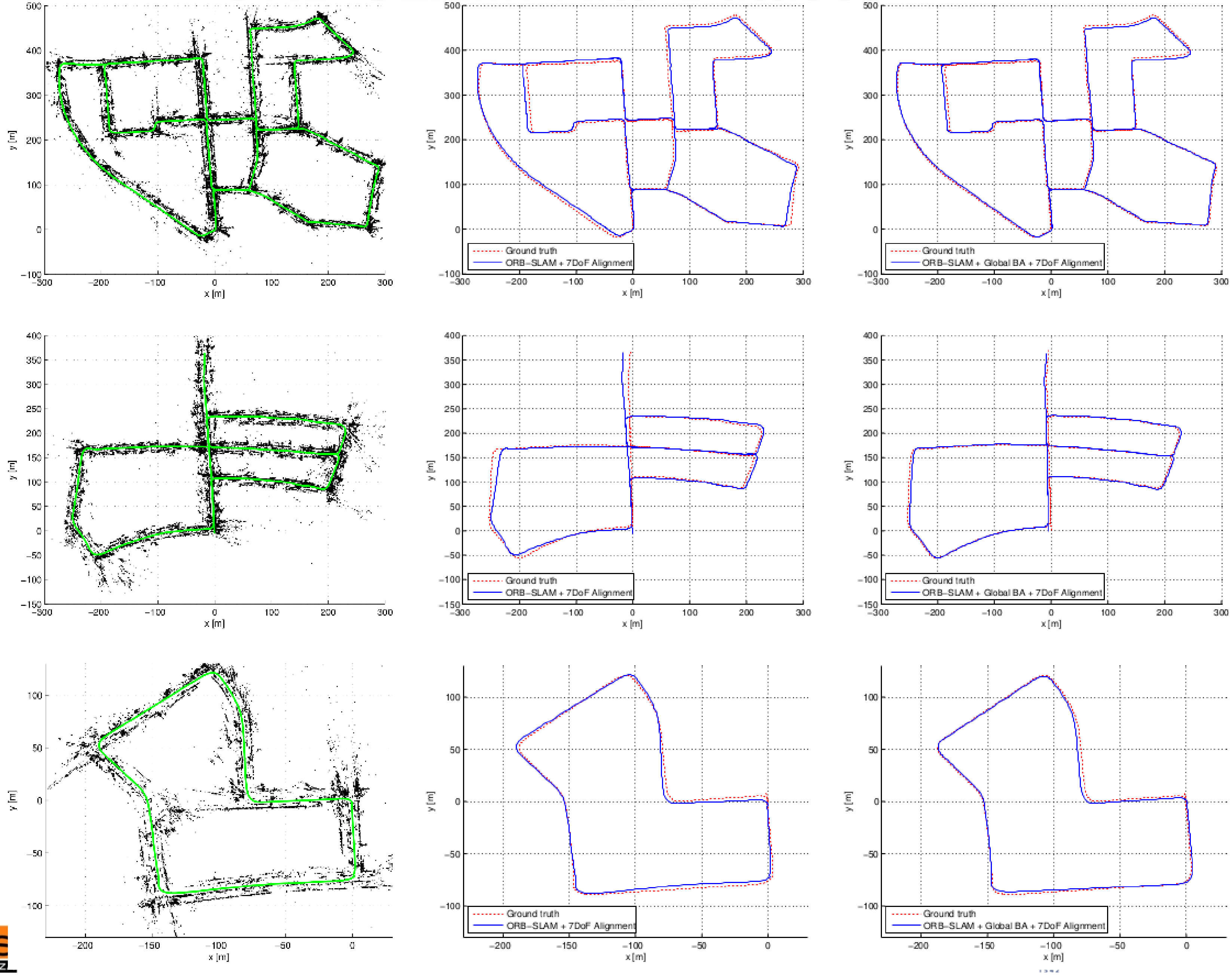
RMSE (cm)

Feature-Based .vs. Direct SLAM

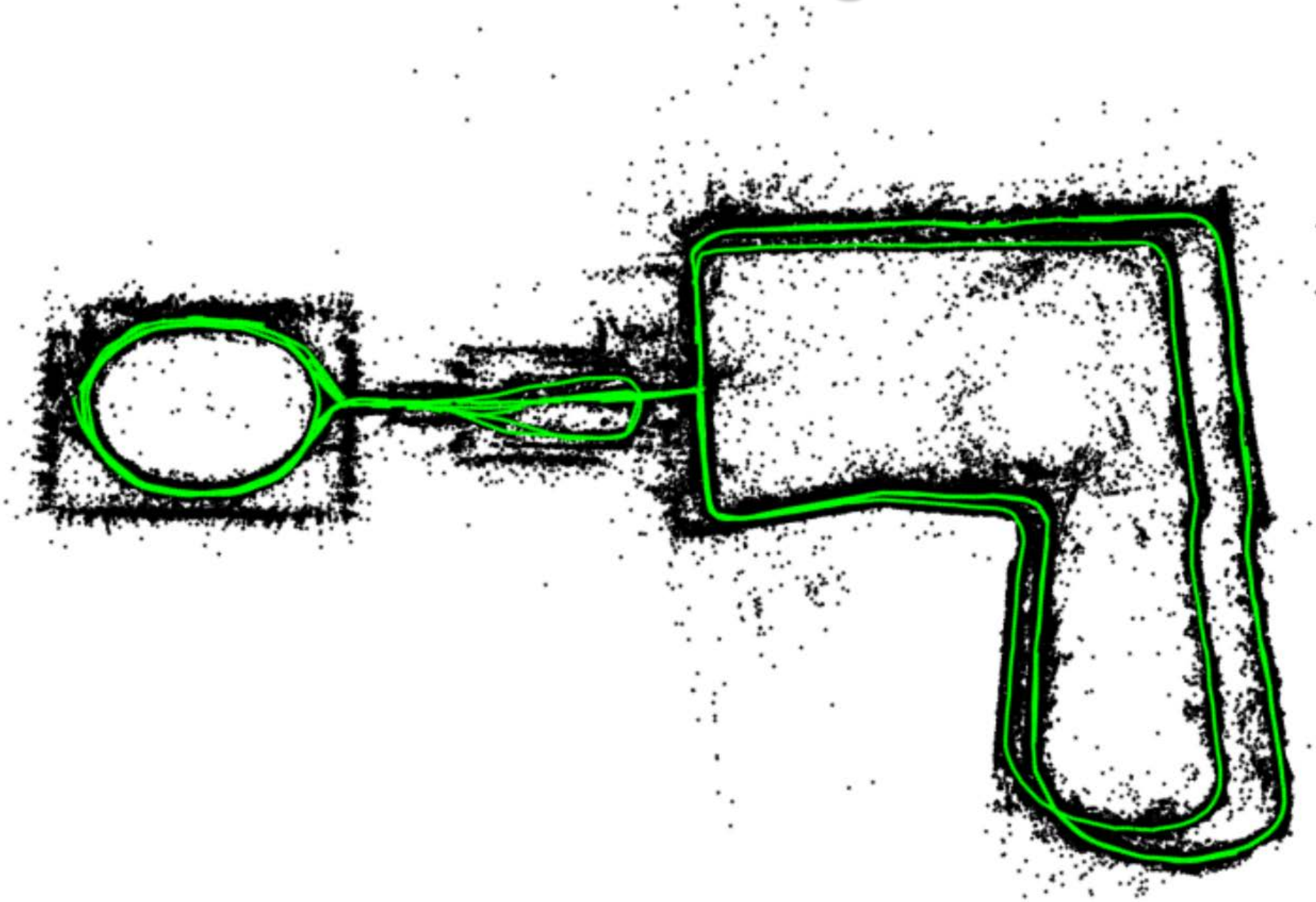
	Feature-Based SLAM (ORB-SLAM)	**Direct SLAM** (LSD-SLAM)
Matching	Illumination/viewpoint invariance *Wide baseline*	Photometric consistency *Narrow baseline*
Map Optimization	Local BA + Global Pose Graph	Global Pose Graph
Loop Detection	Integrated Place Recognition	Need Features (FabMap)
# Points	~300	~100K
Strengths	Excellent accuracy Robust in dynamic scenes Robust initialization	Robust in low texture areas Robust under defocus/motion blur
Weakness	Low texture areas Motion blur	Dynamic Objects Strong viewpoint/illumination changes Rolling-shutter cameras

ORB-SLAM outdoors: Kitti Dataset

ORB-SLAM: Kitti Dataset



ORB-SLAM: New College Dataset



Median Tracking Time per Frame: 30ms
Median Mapping Time per KeyFrame: <400ms

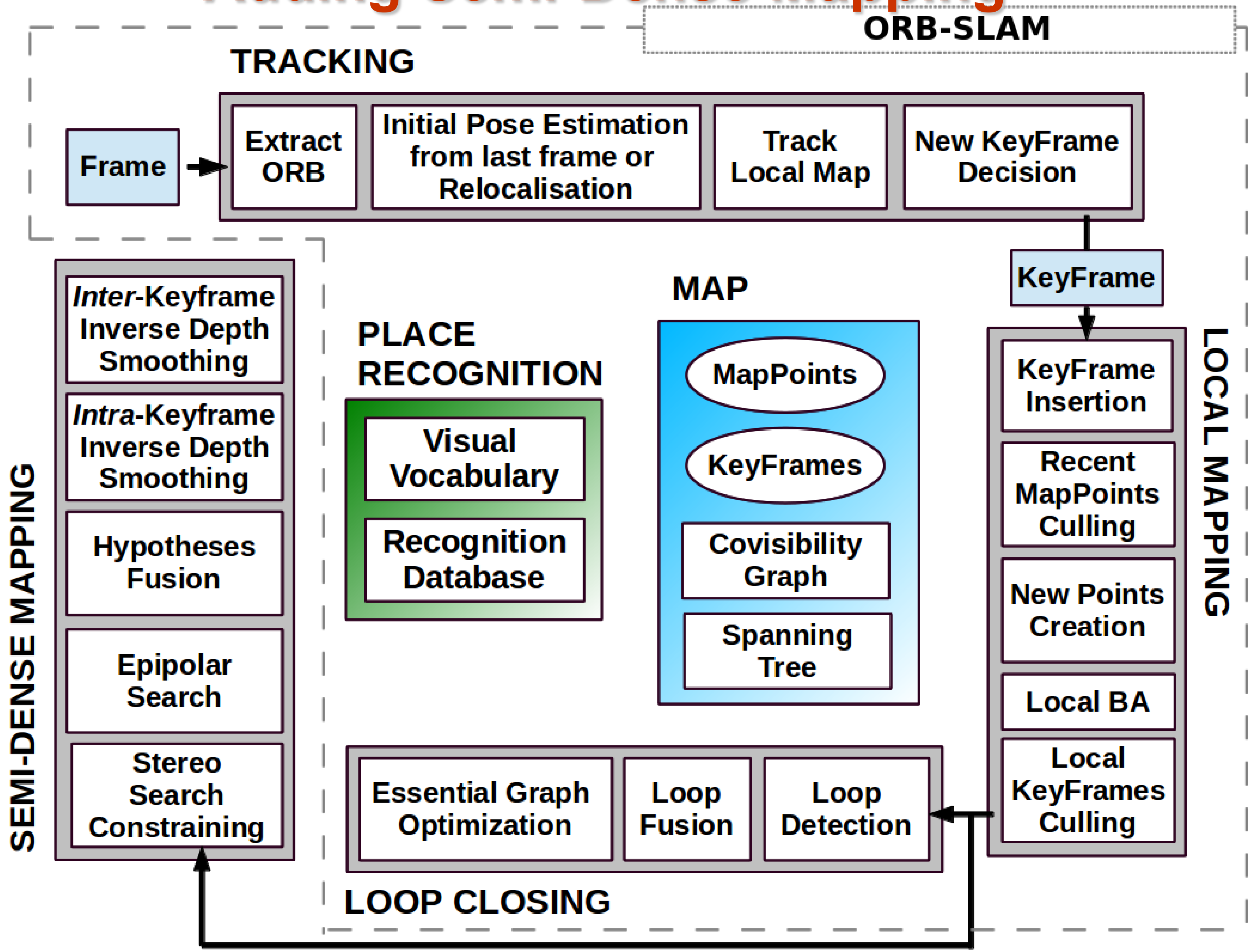
Intel Core i7 @ 2.4 GHz
(4 cores)

Limitations and Extensions

- Limitations
  - Monocular → the absolute scale is unknown
    - » Include in the map a known-size object
  - Requires a reasonably lit area
  - Needs texture: will fail with large plain walls
  - Map is too sparse for interaction with the environment

- Extensions
  - Improve agility using IMU
  - Stereo camera: real scale and robustness to quick motions
  - RGB-D camera: track with features + dense map
  - Semi-dense or dense mapping

Adding Semi-Dense Mapping



ORB-SLAM + Semi-Dense Mapping

Comparison with LSD-SLAM

fr2/desk

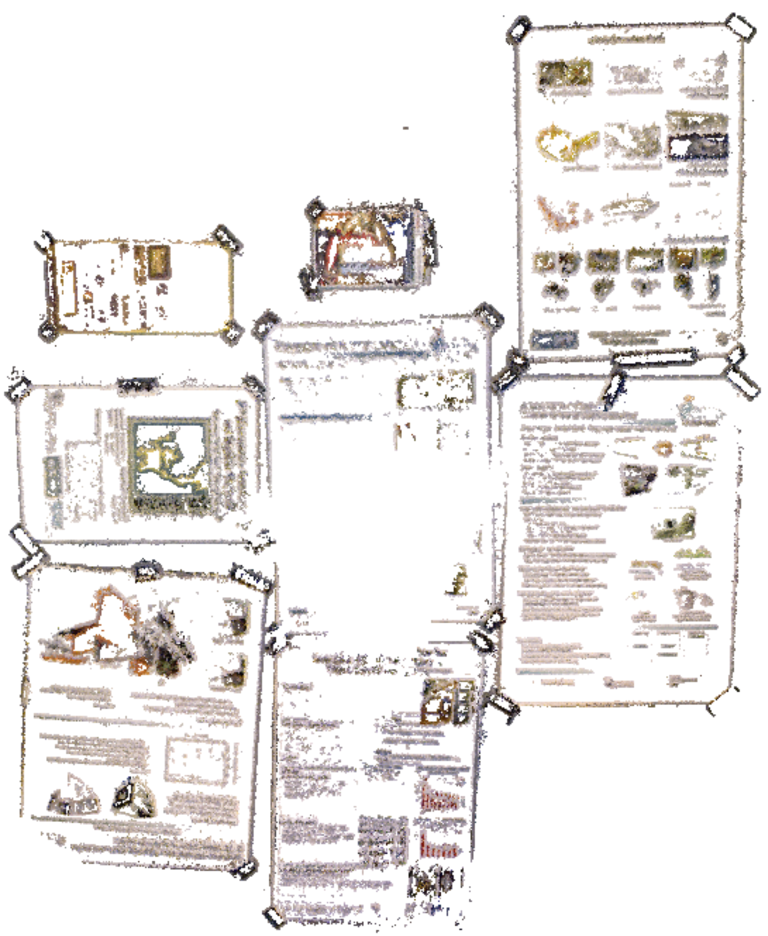
ORB-SLAM

LSD-SLAM

Comparison with LSD-SLAM

fr3/nostructure_texture_near_withloop

ORB-SLAM



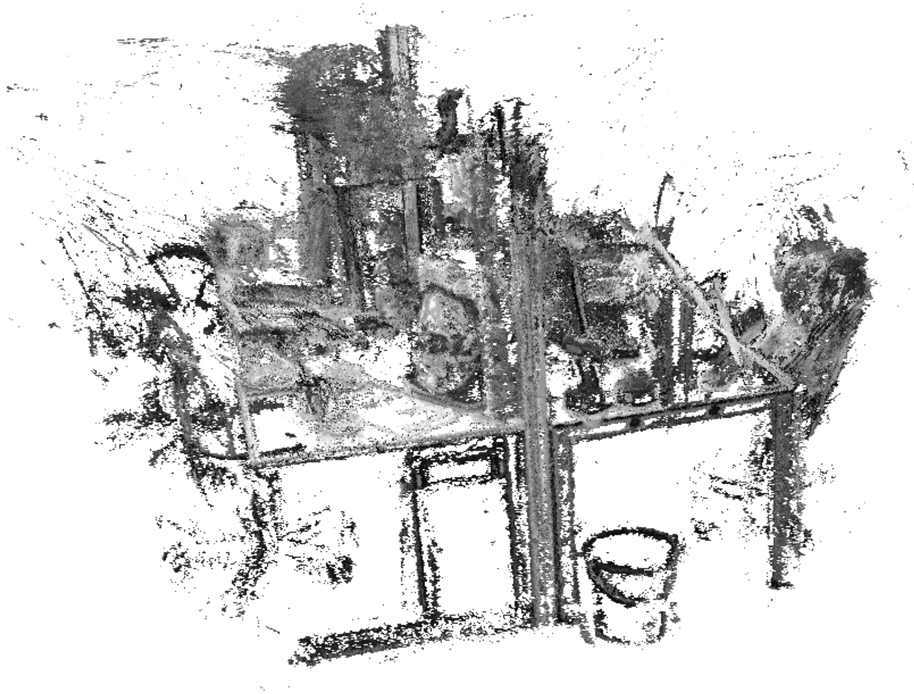
LSD-SLAM

Comparison with LSD-SLAM

fr3/long_office_household

ORB-SLAM

LSD-SLAM



Comparison with LSD-SLAM

fr3/structure_texture_near

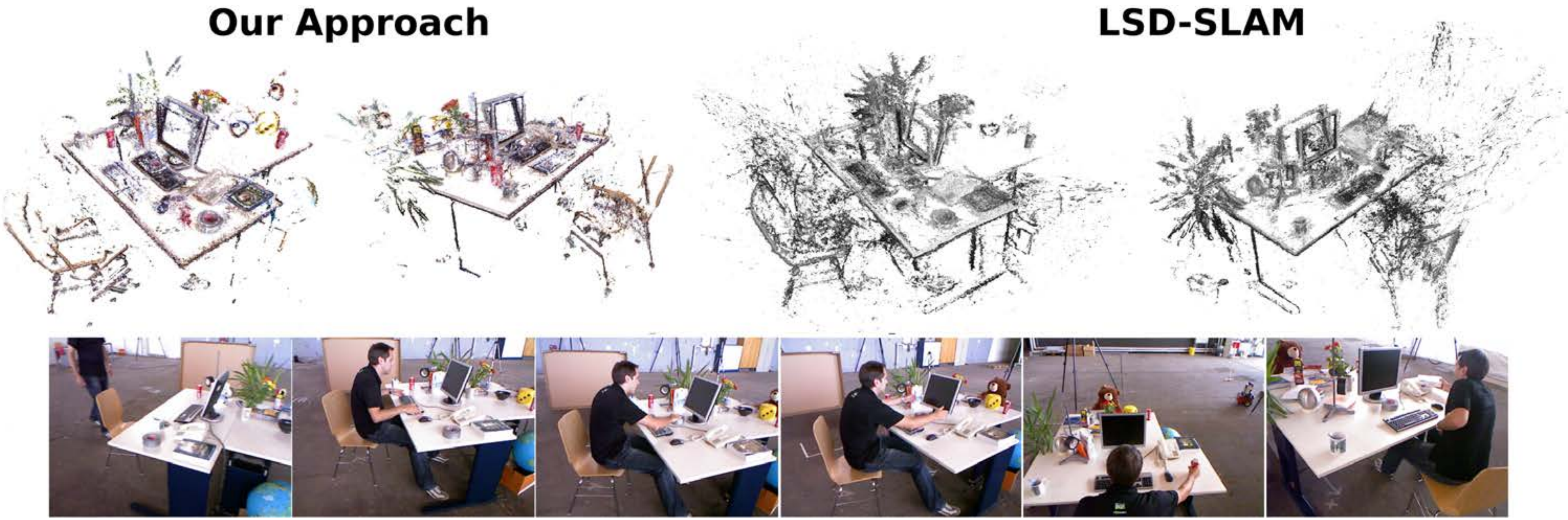
ORB-SLAM

LSD-SLAM

TRACKING FAILURE

Comparison with LSD-SLAM

fr2/desk_with_person



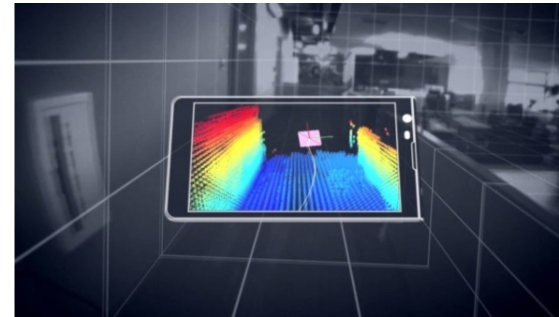
Application to VR

- ORB-SLAM: Feature-Based monocular SLAM
  - Tracking with ORB in <30ms
  - Local mapping: survival of the fittest for keyframes
  - Relocation and loop closing with good viewpoint invariance

- Application to Virtual Reality:
  - Map creation in real-time
  - Viewpoint tracking
    - » Mapping and loop closing can be disabled
    - » Zero-drift tracking
    - » Typical accuracy indoors: 1-3cm
    - » Possible to track multiple users, with different cameras

Some VR Projects using Visual SLAM

- Project Tango (Google)
  - Visual-inertial odometry
  - Area learning (SLAM)
  - RGB-D sensor



- Hololens (Microsoft)
  - RGB-D sensor

- Magic Leap
  - ???

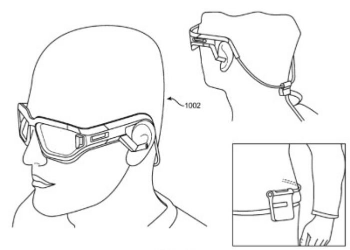
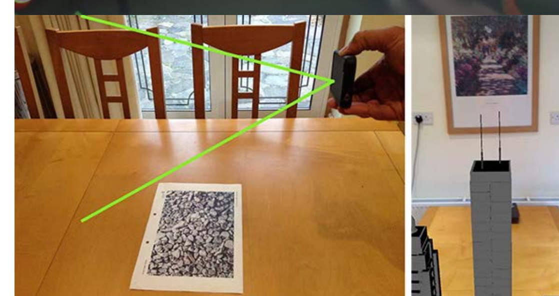


FIG. 16

- Vuforia (Qualcomm)
  - Extended tracking (monocular SLAM)



- Apple (metaio), Oculus (surrealVision),

Monocular SLAM for AR in Medicine

Summary

- ORB-SLAM: Feature-Based monocular SLAM
  - Large-scale operation
  - Loop closing and relocation
  - Robust operation, indoors, outdoors
  - Excellent precision
  - Open-source version released under GPLv3
- More Information:
  - Raúl Mur-Artal and Juan D. Tardós
    Fast Relocalisation and Loop Closing in Keyframe-Based SLAM, ICRA 2014
  - Raúl Mur-Artal, J.M.M. Montiel and Juan D. Tardós
    ORB-SLAM: A Versatile and Accurate Monocular SLAM System, IEEE Transactions on Robotics (under review), arXiv:1502.00956
  - Raúl Mur-Artal and Juan D. Tardós
    Probabilistic Semi-Dense Mapping from Highly Accurate Feature-Based Monocular SLAM, RSS 2015

http://webdiis.unizar.es/~raulmur/orbslam/

Qualcomm Augmented Reality Lecture Series
Vienna - June 11, 2015

ORB-SLAM: a Real-Time Accurate Monocular SLAM System

Juan D. Tardós, Raúl Mur Artal, José M.M. Montiel
Universidad de Zaragoza, Spain
robots.unizar.es/SLAMlab