Visual SLAM

KKY/RVB

Ing. Petr Neduchal

Department of Cybernetics
Faculty of Applied Sciences
University of West Bohemia

ESF projekt Západočeské univerzity v Plzni
reg. č. CZ.02.2.69/0.0/0.0/16_015/0002287

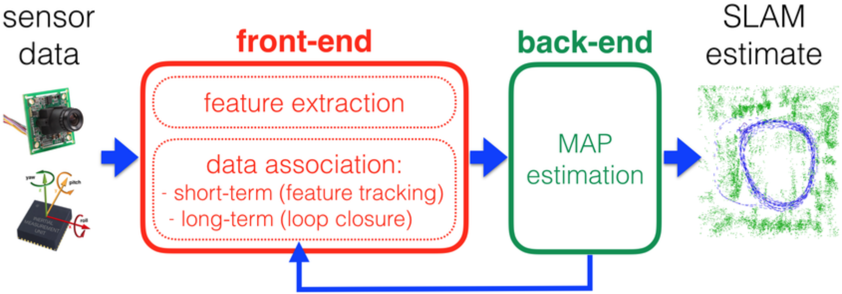
EVROPSKÁ UNIE
Evropské strukturální a investiční fondy
Operační program Výzkum, vývoj a vzdělávání

MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY

DEPARTMENT OF CYBERNETICS

Definition of SLAM

- Problem of building a globally consistent representation of the environment by leveraging both ego-motion compensation and loop-closure (Cadena et al. 2016)
- Estimating the camera trajectory while reconstructing the environment.

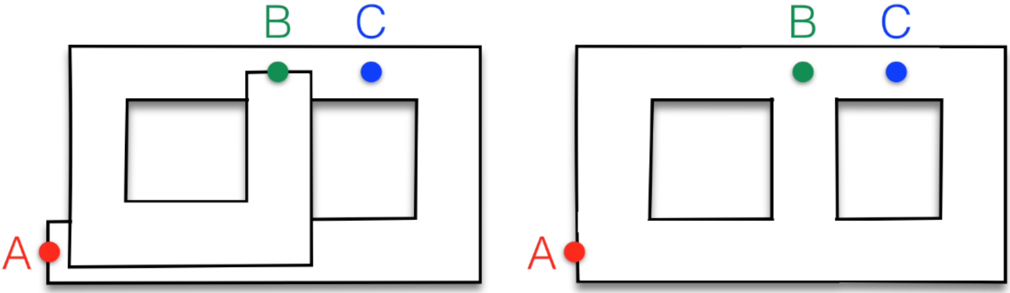


SLAM probability distribution

$$p(x_{0:T}, m \mid z_{1:T}, u_{1:T})$$

distribution path map given observations controls

Visual Odometry vs. Visual SLAM



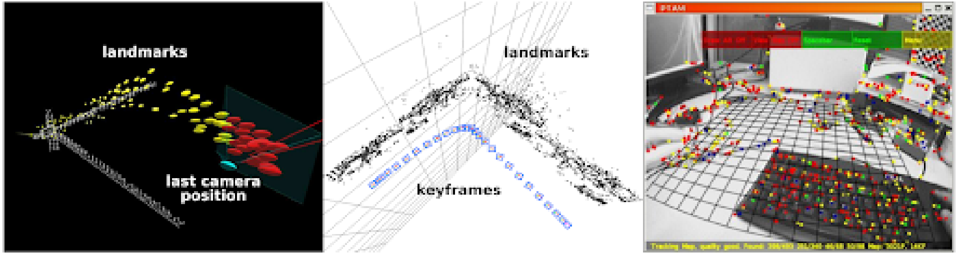
EVROPSKÁ UNIE
Evropské strukturální a investiční fondy
Operační program Výzkum, vývoj a vzdělávání
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY

SLAM examples

- **At home:** vacuum cleaner, lawn mower
- **Air:** surveillance with unmanned air vehicles
- **Underwater:** reef monitoring
- **Underground:** exploration of abandoned mines
- **Space:** terrain mapping for localization

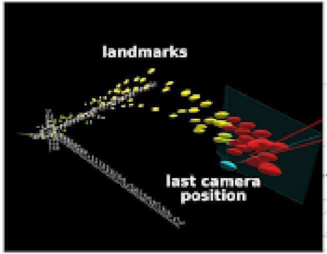
Stone age

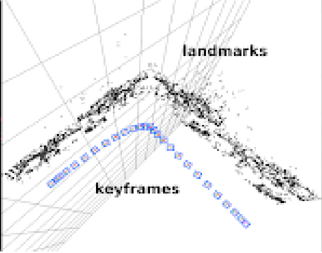
- Early VO works were motivated by NASA Mars exploration program. (90s)
- A. Davison – MonoSLAM (early 2000s)
  - first single camera V-SLAM system
  - using Bayesian Filtering (Extended Kalman Filter)
  - sparse map, full state vector 13 for robot and n for map
  - local image patches to represent landmarks

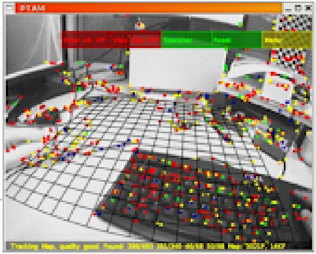


MonoSLAM motion model

$$\hat{x} = \begin{bmatrix} r \\ q \\ v \\ \omega \end{bmatrix}, \qquad f = \begin{bmatrix} r_k + (v_k + V)\Delta t \\ q_k \times Q((\omega_k + \Omega)\Delta t) \\ v_k + V \\ \omega_k + \Omega \end{bmatrix}, \qquad \begin{bmatrix} V \\ \Omega \end{bmatrix} = \begin{bmatrix} \alpha \Delta t \\ \beta \Delta t \end{bmatrix}$$

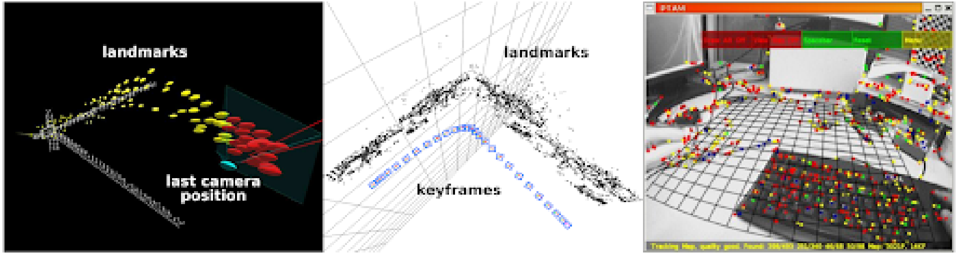






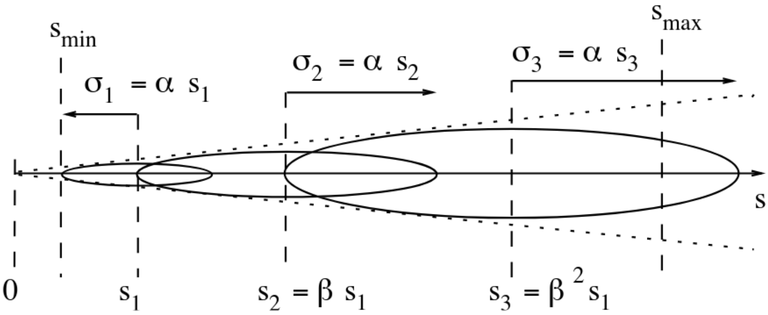
MonoSLAM measurement model

$$h = \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} u_0 - f_u \frac{h_x^R}{h_z^R} \\ v_0 - f_v \frac{h_y^R}{h_z^R} \end{bmatrix}, \qquad h^R = (y - r)]$$



Feature initialization

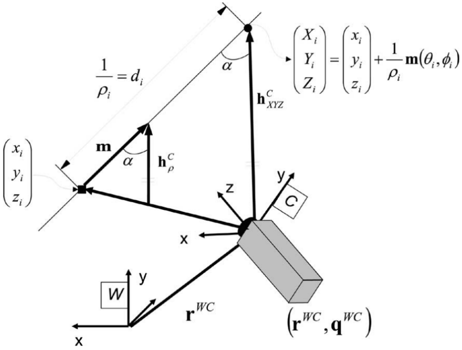
- delayed
- undelayed
  - multiple hypothesis initialization



Feature initialization - Inverse Depth Parametrization

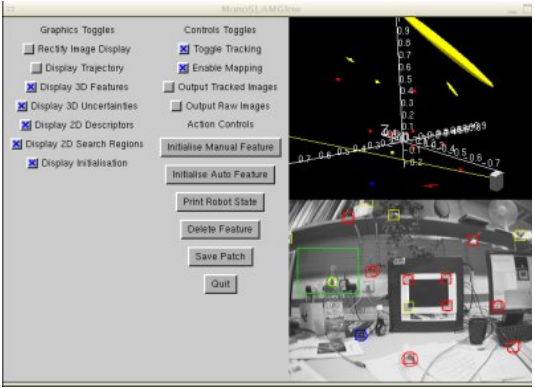
- delayed
- undelayed
  - multiple hypothesis initialization
  - Inverse Depth Parametrization

$$\mathbf{y} = (x_i, y_i, z_i, \theta, \phi, \rho)^T$$
$$m = (cos\phi sin\theta, -sin\phi, cos\phi cos\theta)$$



MonoSLAM and its extensions

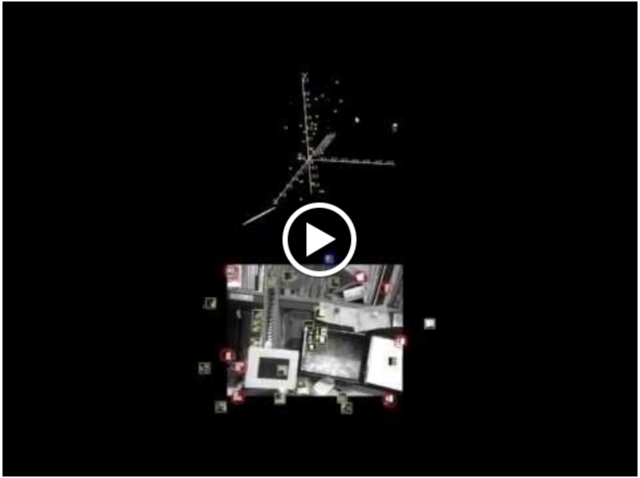
- SceneLib 1 - A. Davison
  - implementation in C++
- SceneLib 2 - Hanme Kim
  - reimplementation
  - C++, modern libraries
- 1-point Ransac for EKF filtering
  - Implementation in MATLAB



EVROPSKÁ UNIE
Evropské strukturální a investiční fondy
Operační program Výzkum, vývoj a vzdělávání
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY

DEPARTMENT OF CYBERNETICS

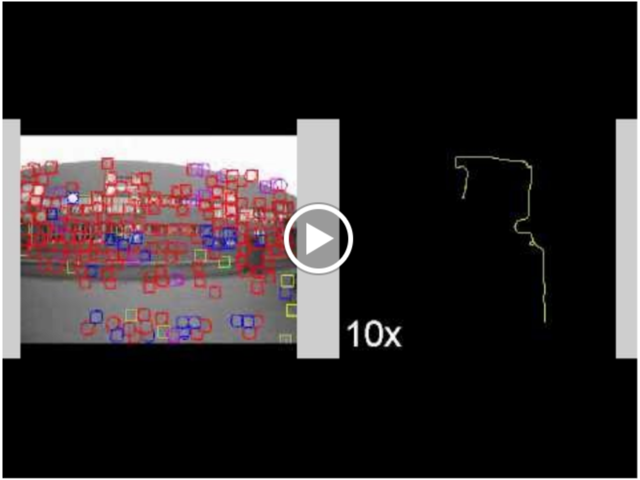
MonoSLAM - Single Camera SLAM



EVROPSKÁ UNIE
Evropské strukturální a investiční fondy
Operační program Výzkum, vývoj a vzdělávání
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY

DEPARTMENT OF CYBERNETICS

1-Point RANSAC for EKF-Based Structure from Motion

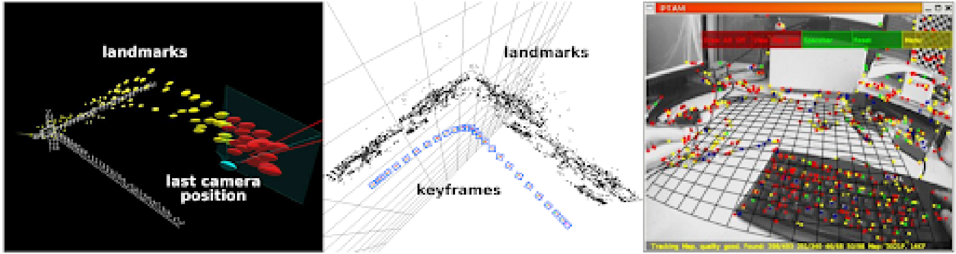


EVROPSKÁ UNIE
Evropské strukturální a investiční fondy
Operační program Výzkum, vývoj a vzdělávání
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY

DEPARTMENT OF CYBERNETICS

PTAM - Klein, Murray 2007

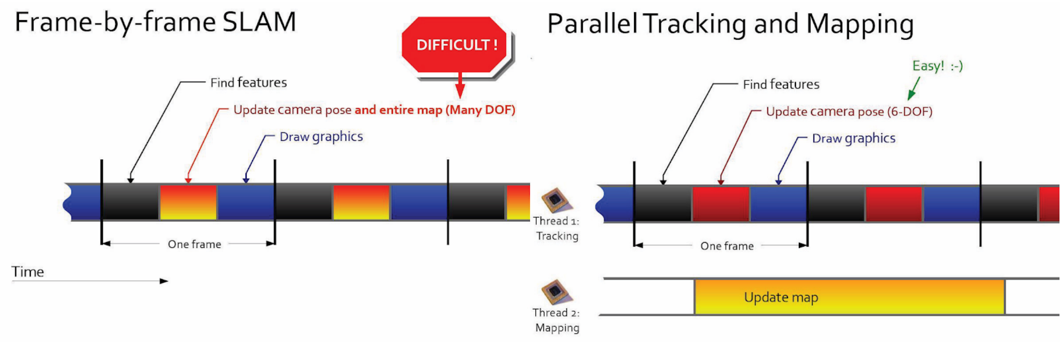
- feature based SLAM algorithm
- parallelizing the tracking and mapping tasks
- keyframe-based Bundle Adjustment instead of filtering – Graph SLAM
- designed for small scale AR application



EVROPSKÁ UNIE
Evropské strukturální a investiční fondy
Operační program Výzkum, vývoj a vzdělávání
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY

DEPARTMENT OF CYBERNETICS

PTAM - Threads



EVROPSKÁ UNIE
Evropské strukturální a investiční fondy
Operační program Výzkum, vývoj a vzdělávání
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY

PTAM: The Map

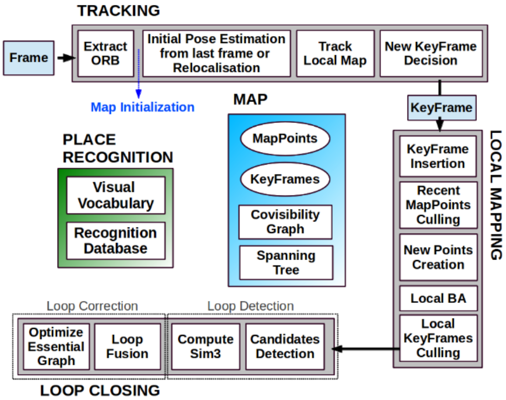
- The map consists of point features (locally planar patches)
- It contains keyframes (snapshots) – i.e. frame that is selected as a represent for a set of consecutive frames
- Each keyframe stores 4-level pyramid of grayscale images
- Point feature is stored with:
  - reference to a keyframe (first observation)
  - pyramid level
  - pixel location

Parallel Tracking and Mapping for Small AR Workspaces (PTAM)

Parallel Tracking and Mapping for Small AR Workspaces (PTAM) - extra

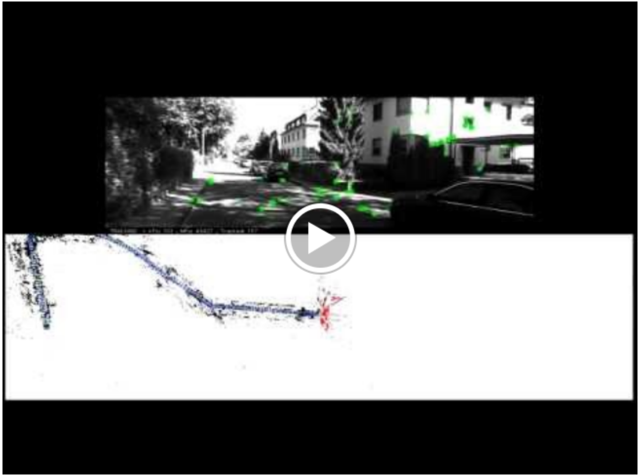
ORB-SLAM 1 and 2 (2014-2016)

- State of The Art
- feature based
- three threads
  - tracking
  - local mapping
  - global mapping
  - ORB features used in all threads
- Bag Of Visual Words used for relocalization.



EVROPSKÁ UNIE
Evropské strukturální a investiční fondy
Operační program Výzkum, vývoj a vzdělávání
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY

ORB-SLAM in the KITTI dataset (Sequence 00)



EVROPSKÁ UNIE
Evropské strukturální a investiční fondy
Operační program Výzkum, vývoj a vzdělávání
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY

DEPARTMENT OF CYBERNETICS

ORB-SLAM2: an Open-Source SLAM for Mono, Stereo and RGB-D

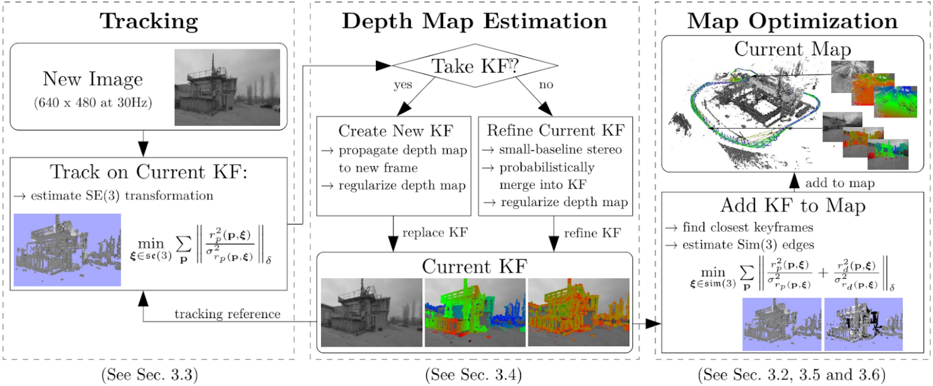
DTAM: Dense Tracking and Mapping in Real-Time

- photometric error instead of reprojection error
- all pixels are processed – whole image alignment
- Computationaly expensive - GPU

EVROPSKÁ UNIE
Evropské strukturální a investiční fondy
Operační program Výzkum, vývoj a vzdělávání
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY

LSD SLAM, DSO

- semi-dense - only pixels with non-negligible gradient
- depth map propagated from frame to frame – update
- Gauss-Newton Optimization

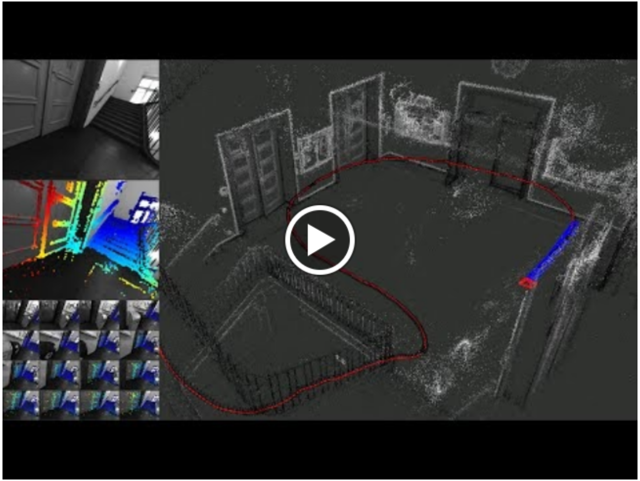


LSD-SLAM: Large-Scale Direct Monocular SLAM (ECCV '14)

EVROPSKÁ UNIE
Evropské strukturální a investiční fondy
Operační program Výzkum, vývoj a vzdělávání
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY

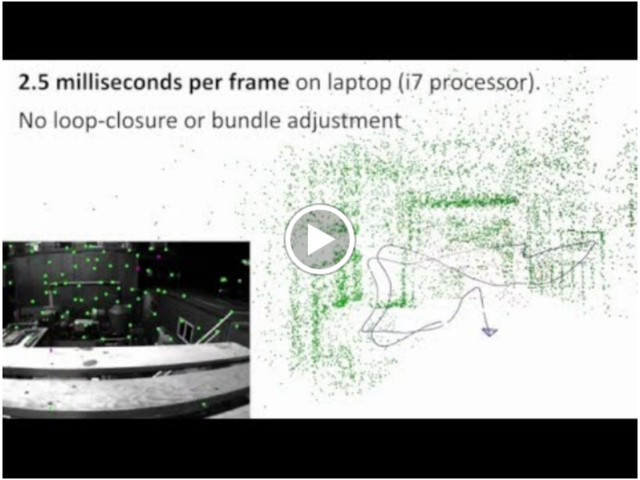
DEPARTMENT OF CYBERNETICS

DSO: Direct Sparse Odometry

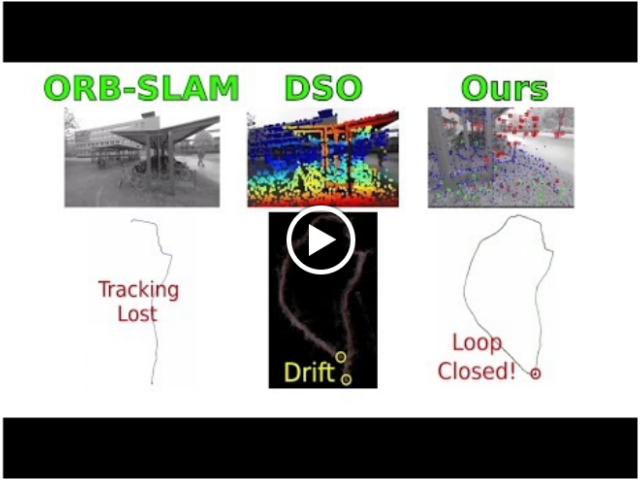


SVO: Fast Semi-Direct - Hybrid system

- image Alignment
- features for initialization of new 3D points
- Loosely-Coupled Semi-Direct Monocular SLAM – SVO with Loop Closure



Loosely-Coupled Semi-Direct Monocular SLAM



EVROPSKÁ UNIE
Evropské strukturální a investiční fondy
Operační program Výzkum, vývoj a vzdělávání
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY

Current Research – 2018

- neural networks
- Visual Inertial SLAM
- CVPR 2018 workshop – Deep Learning for Visual SLAM

Current Research

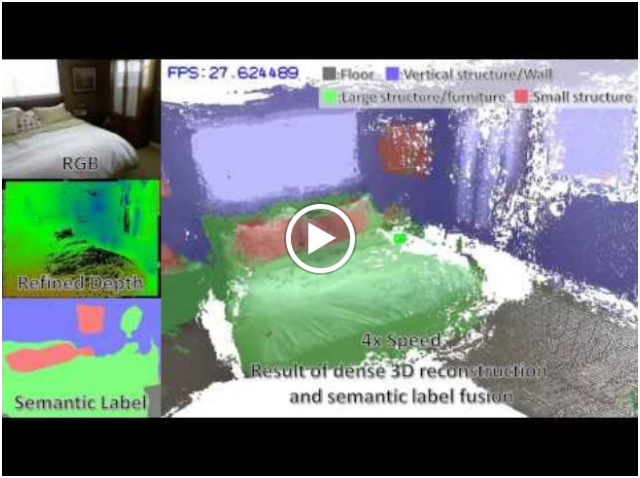
- SuperPoint: Self-Supervised Interest Point Detection and Description
- Global Pose Estimation With an Attention-Based Recurrent Network
- Visual SLAM for Automated Driving: Exploring the Applications of Deep Learning
- Mask-SLAM: Robust Feature-Based Monocular SLAM by Masking Using Semantic Segmentation
- Geometric Consistency for Self-Supervised End-to-End Visual Odometry
- DepthNet: A Recurrent Neural Network Architecture for Monocular Depth Prediction
- Monocular Depth Prediction Using Generative Adversarial Networks
- Learning 3D Scene Semantics and Structure From a Single Depth Image
- Keyframe-Based Visual-Inertial SLAM Using Nonlinear Optimization
- Unsupervised Learning of Depth and Ego-Motion from Video

EVROPSKÁ UNIE
Evropské strukturální a investiční fondy
Operační program Výzkum, vývoj a vzdělávání
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY

DEPARTMENT OF CYBERNETICS

OKVIS: Open Keyframe-based Visual-Inertial SLAM

CNN-SLAM: Real-time dense monocular SLAM with learned depth prediction



Thank you for your attention!

Questions?

EVROPSKÁ UNIE
Evropské strukturální a investiční fondy
Operační program Výzkum, vývoj a vzdělávání

MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY

DEPARTMENT OF CYBERNETICS