



HAL
open science

Dense, metric and real-time 3D reconstruction for autonomous drone navigation

Yassine Habib, Panagiotis Papadakis, Cédric Buche, Cédric Le Barz, Antoine Fagette

► **To cite this version:**

Yassine Habib, Panagiotis Papadakis, Cédric Buche, Cédric Le Barz, Antoine Fagette. Dense, metric and real-time 3D reconstruction for autonomous drone navigation. Journées des Jeunes Chercheurs en Robotique (JJCR 2021), Oct 2021, Paris, France. 2021. hal-04184995

HAL Id: hal-04184995

<https://hal.science/hal-04184995>

Submitted on 22 Aug 2023

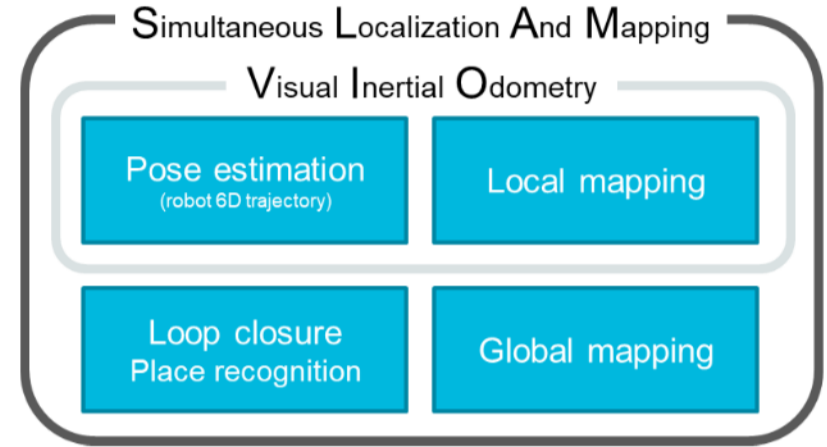
HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1. Context:

- ▶ Autonomous navigation
 - **Dense** and **metric** mapping
 - **Real-time** application on aerial robots
- ▶ Constraints and challenges
 - Limited computation resources, low energy consumption
 - Use of passive sensors only
 - Restriction to monocular cameras
 - Fast motions
 - Illumination changes

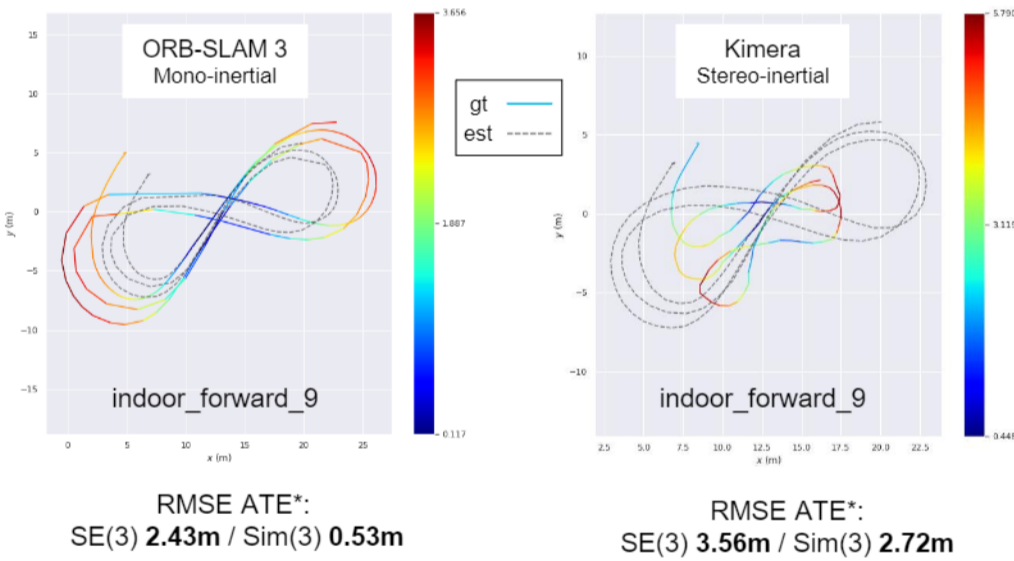
2. Approach:



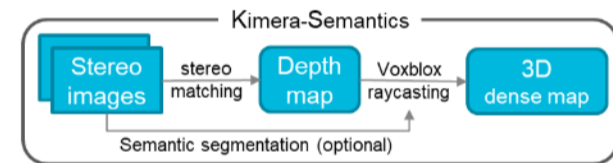
- ▶ State-of-the-art methods:
 - **ORB-SLAM3** [2] (Mono/Stereo inertial, RGB-D)
 - **Kimera** [1] (Stereo inertial)

3. Indicative performance of state-of-the-art

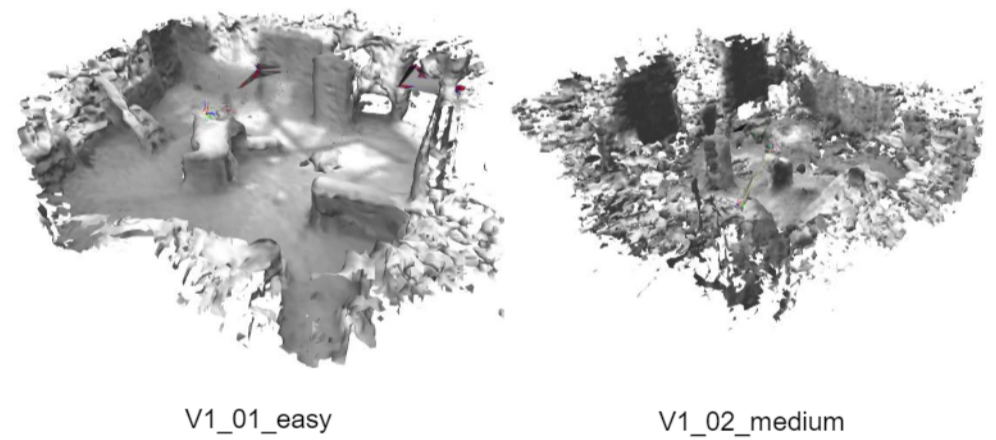
VIO evaluation on UZH FPV Drone Racing dataset [3]



* Root Mean Square Error (RMSE) of the Absolute Trajectory Error (ATE)



Kimera 3D reconstruction on EuRoC [4] dataset



4. Monocular depth estimation

- ▶ As we restrain our study to monocular cameras, stereo-matching is replaced by Deep Learning-based depth estimation.

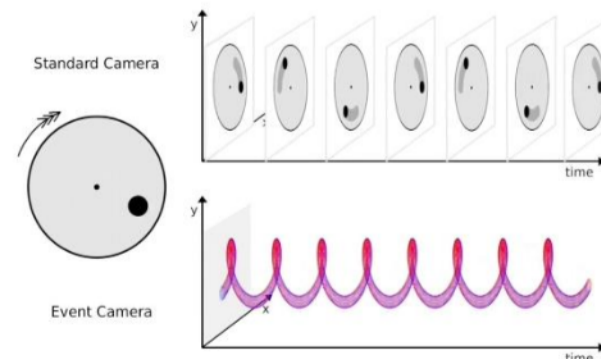


PackNet [5] monocular depth map prediction

| | |
|---------------|--|
| Training | Supervised, Unsupervised, Semi-supervised |
| Architectures | CNN, Auto-encoder, RNN, GAN |
| Task | Single task: Depth map Multi-tasks: Depth map, visual odometry, optical flow, semantic segmentation |

5. Event cameras

- ▶ The characteristics of the sensor are particularly suitable for our context. It is robust to fast motions and high illumination changes.



| | |
|--|--------------------------------|
| High Dynamic Range: 120 dB | No direct intensity value |
| Asynchronous High update rate: 1MHz | No data in static scenes |
| Low Energy consumption: 20 mW | Few computer vision algorithms |

[1] Rosinol, Antoni et al. "Kimera: an Open-Source Library for Real-Time Metric-Semantic Localization and Mapping." 2020 IEEE International Conference on Robotics and Automation (ICRA).

[2] Campos, Carlos et al. "ORB-SLAM3: An Accurate Open-Source Library for Visual, Visual-Inertial, and Multimap SLAM". IEEE Transactions on Robotics. (2021): 1-17.

[3] Delmerico, Jeffrey et al. "Are We Ready for Autonomous Drone Racing? The UZH-FPV Drone Racing Dataset." 2019 International Conference on Robotics and Automation (ICRA).

[4] Burri, Michael et al. "The EuRoC micro aerial vehicle datasets". The International Journal of Robotics Research. (2016).

[5] Guizilini, Vitor et al. "3D Packing for Self-Supervised Monocular Depth Estimation." 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)