Avancement thèse

César D. & Damien V.

CENTRE NATIONAL D'ÉTUDES SPATIALES

Institut Supérieur de l'Aéronautique et de l'Espace

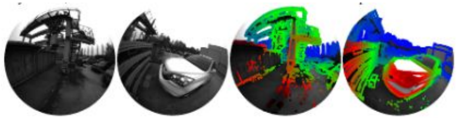
SUPAERO

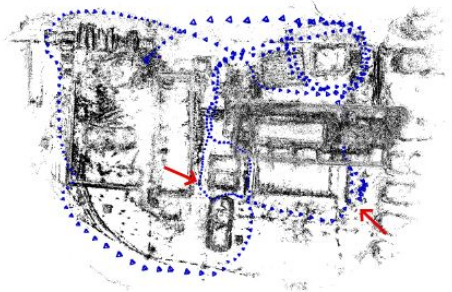
Biblio : SLAM fisheye

-> VINS-MONO: propose un modèle fisheye mais qui est converti en pinhole

-> LSD Omni: SLAM dense mono fisheye qui propose une étude de différents modèles caméra

-> ORB-SLAM 3: propose une classe abstraite caméra pour tester d’autres modèles, mais n’a fait aucune expériences sur un banc stéréo fisheye





Toujours pas d’expériences stereo fisheye pour du SLAM éparse dans la littérature

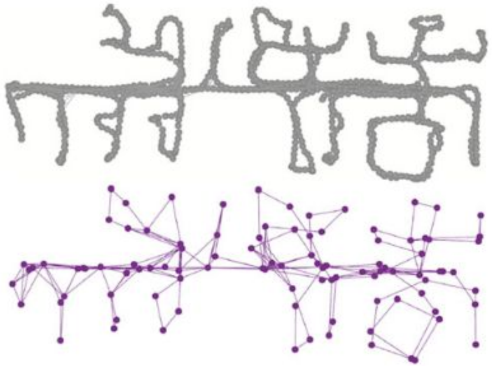
Biblio: Backend SLAM

-> efficacité de l’optimiseur (ISAM2, g2o)

-> marginalization et sparsification du graph (J. Vallvé)

-> problématiques qui se pose pour un SLAM à grande échelle (Manhattan 3500…)

-> Séjour à Barcelone (IRI) du 28/03 au 8/04 pour travailler sur du graph SLAM visuel



Soumission IROS 2022: CosySLAM

-> SLAM visuel inertiel basé deep learning avec un backend factor graph

-> Expériences sur dataset public avec comparaison ORB-SLAM 3 et VINS MONO

-> études des performances en terme de temps de calcul

-> expériences de robustesses

CosySlam: investigating object-level SLAM for detecting locomotion surfaces

César Debeunne[‡†], Médéric Fourmy[†], Yann Labbé[△], Pierre-Alexandre Léziart[†], Guilhem Saurel[†], Joan Solà[†*] and Nicolas Mansard[†+]

Abstract— **While blindfolded legged locomotion has demonstrated impressive capabilities in the last few years, further progresses are expected from using exteroceptive perception to better adapt the robot behavior to the available surfaces of contact. In this paper, we investigate whether mono cameras are suitable sensors for that aim. We propose to rely on object-level SLAM, fusing RGB images and inertial measurements, to simultaneously estimate the robot balance state (orientation in the gravity field and velocity), the robot position, and the location of candidate contact surfaces. We used CosyPose, a learning-based object pose estimator for which we propose an empirical uncertainty model, as the sole front-end of our visual inertial SLAM. We then combine it with inertial measurements which ideally complete the system observability, although extending the proposed approach would be straightforward (e.g. kinematic information about the contact, or a feature based visual front end). We demonstrate the interest of object-based SLAM on several locomotion sequences, by some absolute metrics and in comparison with other mono SLAM.**

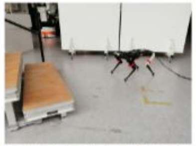


Fig. 1: Experimental setup: a RealSense D435i is mounted on the Solo robot which localizes itself with respect to stairs. A motion capture system provides ground truth of the robot pose.

I. INTRODUCTION

State estimation is a central aspect of the design of legged robots systems. Using estimates of the base velocity and gravity direction, modern locomotion controllers [1], [2] can achieve remarkable robustness without precise knowledge of their environment. In these cases, the state estimation relies simply on inertial, kinematic, and contact measurements to achieve balance [3]–[5]. However, to accomplish reactive contact planning over uneven surfaces [6], we need to go beyond a blindfolded system by using exteroceptive sensors.

Most of the current approaches exploit depth sensors which can be used to build local elevation maps of the robot surroundings [7] later processed to extract the contact surfaces needed by locomotion planners [8]. Current methods loosely couple the elevation mapping with the estimation of the robot state. Other approaches strive to extend reconstruction capabilities by using 3D voxels [9], dense surfaces [10], or meshes representations of the environment [11].

‡ ISAE-Supaero, Toulouse
† LAAS-CNRS, Université de Toulouse, France
△ Inria, École normale supérieure, CNRS, PSL Research University, Paris, France
* Institut de Robòtica i Informàtica Industrial, Barcelona
+ Artificial and Natural Intelligence Toulouse Institute, Toulouse, France

This work has been supported by the MEMMO European Union project within the H2020 Program under Grant Agreement No. 780684, by the HPC resources from GENCI-IDRIS (Grant AD011011181R1), the Spanish Ministry of Science, Innovation, and Universities project EB-SLAM (DPI2017-89564-P), by the EU H2020 project GAUSS (DPI2017-89564-P) and by the Spanish State Research Agency through the Maria de Maeztu Seal of Excellence to IRI MDM-2016-0656. The experiments have been performed on the cluster DeepRAP, graciously provided by Patrick Danès (LAAS-CNRS), who we thank respectfully.

On the opposite, recent legged estimators, based on the tight coupling of exteroceptive inertial SLAM system with leg kinematics, have been shown to provide accurate and robust source odometry [12], [13]. The use of sparse features in these approaches is however not intended for contact surfaces extractions. Yet they tend to show the importance of tightly coupling all locomotion estimators, to which we also intend to contribute.

In indoor environments, a localization system may benefit from the presence of known objects. SLAM++ [14] showed that object-level semantic SLAM could be achieved in office environments using depth sensors and scanned models of objects such as chairs and tables. More recent works on the matter rely mainly on deep-learning-based object segmentation to detect objects in RGB images [15]. This information is used along with depth measurements to perform semantic SLAM with static [16] or moving [17]–[19] objects. Using solely monocular cameras, object SLAM can also be achieved without object shape priors by directly integrating segmentation bounding box in classical feature-based SLAM frameworks [20], [21]. Aside from the camera trajectory and object poses, object shapes may also be optimized using a differentiable rendering engine [22].

Object-level SLAM has not been explored in the context of legged robot locomotion yet. In this paper, we propose to build a visual-inertial object-level SLAM by merging an object pose estimator based on shape priors with pre-integrated inertial measurements. We rely on the open-source pose estimator CosyPose [23], for which we first propose

Un bon entraînement à la soumission d’un papier de SLAM

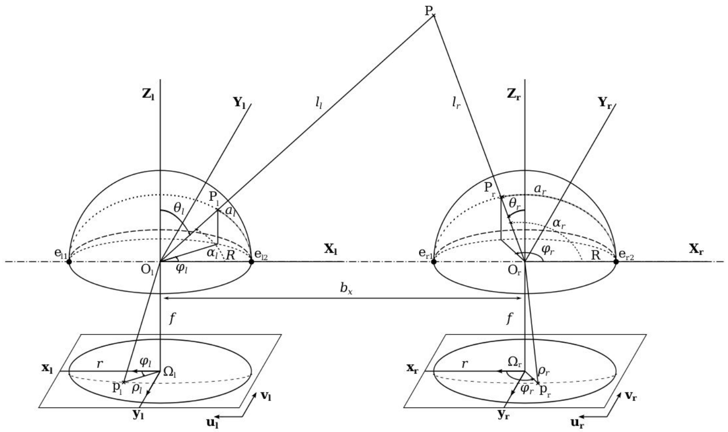
Code: Géométrie épipolaire pour caméras sphériques

-> Estimation of omnidirectional camera model from epipolar geometry - Micusik

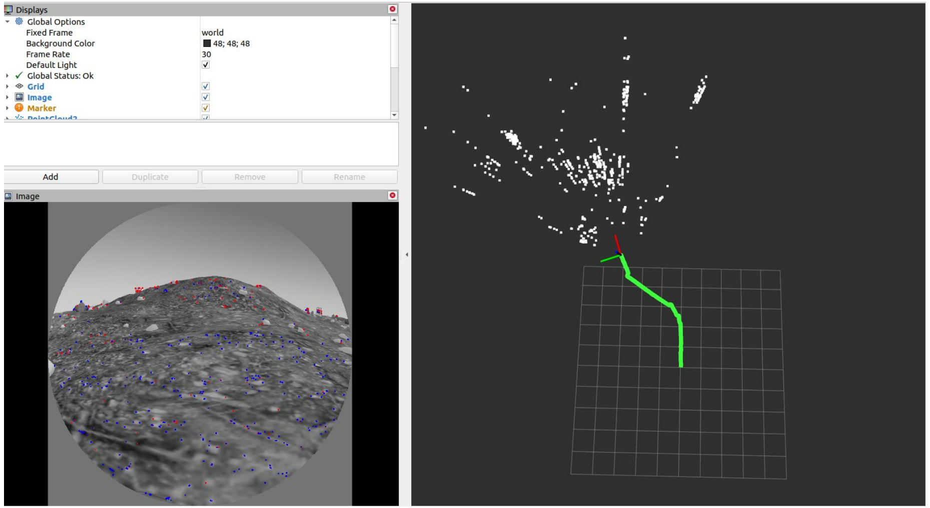
-> Algorithme des 8 points généralisable à tous les modèles de caméra

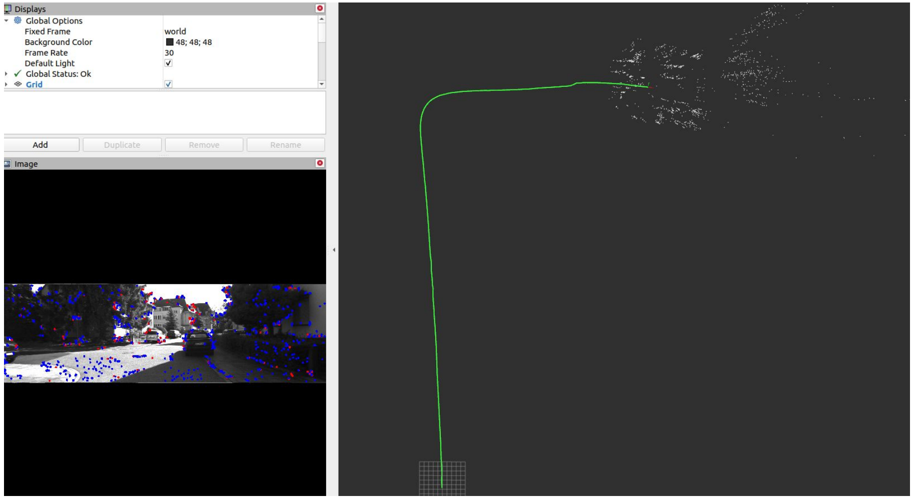
-> Validation du modèle avec le plan épipolaire

$$
\begin{aligned}
q_2^T E_{21} q_1 &= 0 \\
E_{21} &= [{}_1t^2]_x \, {}_1R^2
\end{aligned}
$$



Code: Adaptation de PAVO aux caméras omni





-> La géométrie épipolaire ne donne pas de bons résultats en stationnaire: utiliser une homographie RANSAC / généraliser l’algorithme pnp

-> L’échelle n’est pas estimée: utiliser la stéréo

-> Temps de calcul élevé: tester différents keypoints / descripteurs

Expé: acquisition dans une grotte & claie de portage

FISA/ISAE-SUPAERO 2A — 2022-2023
”Avionique et Systèmes Embarqués”

Intégration et mise au point d’un système d’acquisition de données portatif pour l’exploration d’environnements sous-terrain

Keywords: système d’acquisition mobile, exploration, robotique, navigation

(a) Sac à dos Apple

(b) Sac à dos Google

Figure 1: Exemple de systèmes d’acquisition portables.

Questions toujours ouvertes

-> Comment assurer l’éclairage du FOV? (Quel intérêt d’utiliser des fisheye sur l’éclairage est focalisé sur une partie de l’image)

-> Est-il possible de faire de la fermeture de boucle à bas coup de calcul? (nécessite de la biblio)