Cartographie 3D dense et temps-réel à partir d'un système de vision monoculaire pour l'archéologie sous-marine

Maxime Ferrera

ONERA, The French Aerospace Lab

Gdr Robotique - Robotique, Patrimoine et Archéologie

06 février 2020

Context

ONERA
THE FRENCH AEROSPACE LAB

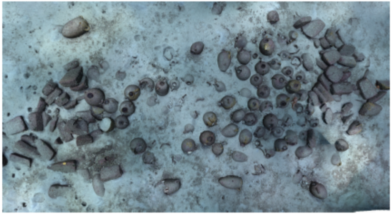
LIRMM

- PhD student from Nov. 2016 - Dec. 2019 at ONERA and LIRMM
- PhD thesis topic : Vision-based Localization and 3D Mapping for Underwater Environments
- Supervisors : Dr. Julien Moras, Dr. Pauline Trouvé-Peloux (ONERA) and Dr. Vincent Creuze (LIRMM)

Context

Underwater Archaeology

- Many sites below 100 meters deep
- Not human-friendly environnements

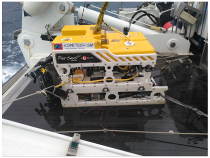


Credit : DRASSM

Context

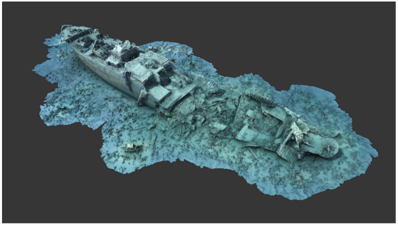
Robots to the rescue

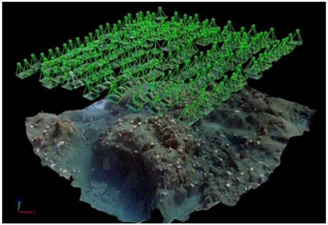
- ROV : Remotely Operated Vehicles
- ROVs are used for deep surveys



Context

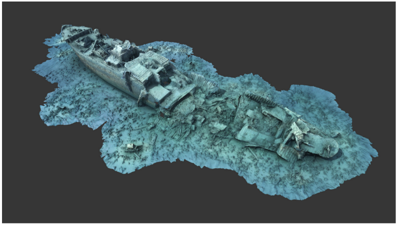
Underwater Photogrammetry

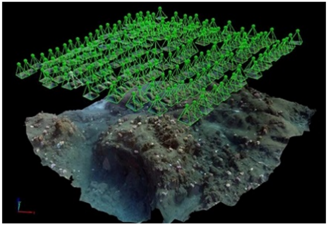




Context

Underwater Photogrammetry

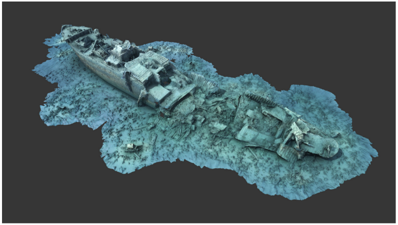


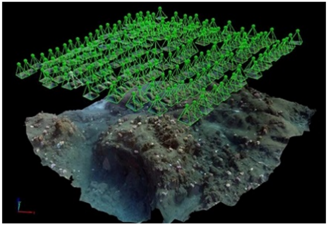


- Softwares : Photoscan, MicMac, Colmap, ...

Context

Underwater Photogrammetry



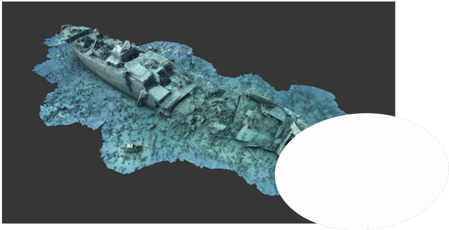


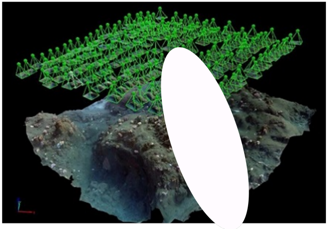
- Softwares : Photoscan, MicMac, Colmap, ...

- Heavy computation $\rightarrow$ offline processing
- Results obtained days to weeks after the mission

Context

Underwater Photogrammetry





- Softwares : Photoscan, MicMac, Colmap, ...

- Heavy computation → offline processing
- Results obtained days to weeks after the mission

Objectives

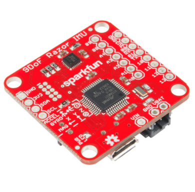
- **Estimate accurate localization in real-time**
- **Produce dense 3D reconstructions**
- **Use low-cost sensors**

Thesis proposal

SLAM from a Monocular Vision-based System

Camera

- Monochromatic Camera
- 20 Hz



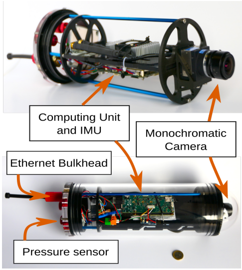
MEMS-IMU

- Inertial meas.
- 200 Hz
- High drift

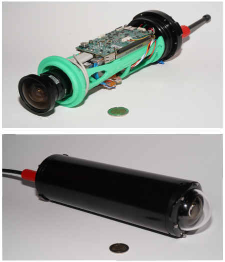
Pressure Sensor

- Pressure meas. (bar) $\propto$ Depth (m)
- 5-10 Hz

Designed Systems



Size : 33.4 $\times$ 11.4 cm
Depth rated : 100 m



Size : 25.8 $\times$ 8.9 cm
Depth rated : 500 m

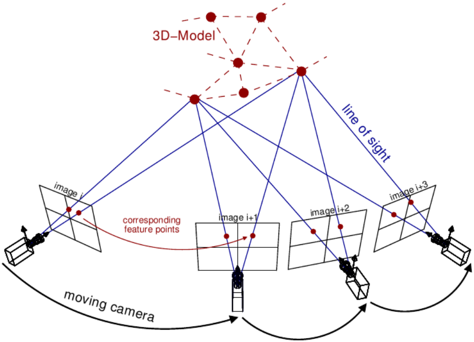
- **On-board computation**
  $\rightarrow$ Autonomous and independent
  $\rightarrow$ No bandwidth issue

- **Compact**
- **Low-cost :** $<$ 2.5 k€

1. **Underwater Monocular Visual SLAM**
2. Multi-Sensor SLAM
3. Monocular Dense 3D Mapping

Underwater Monocular Visual SLAM

SLAM by Structure-from-Motion



From pixel correspondences :

- Localization → 3D map
- 3D map → Localization

Underwater Monocular Visual SLAM

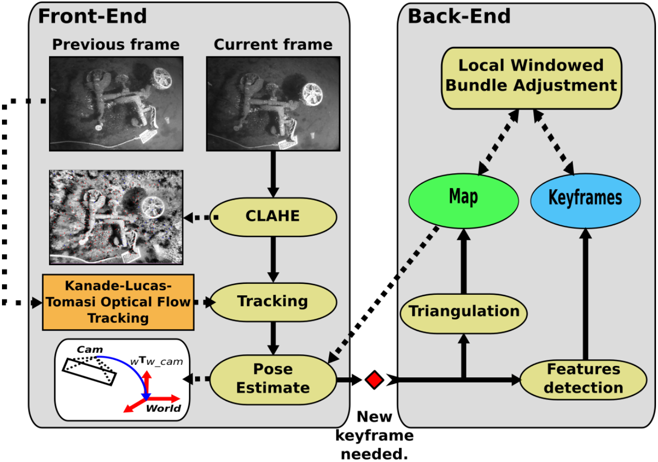
Problem Statement

- Estimate the pose of the camera at each new image
- Pose : $\mathbf{X_i} = (\mathbf{R}, \mathbf{t}) \in \mathbb{SE}(3) \quad | \quad \mathbf{R} \in \mathbb{SO}(3) \quad \mathbf{t} \in \mathbb{R}^3$

- Estimate the position of 3D landmarks : $\mathbf{lm_j} \in \mathbb{R}^3$

Underwater Monocular Visual SLAM

Tracking / Mapping : Two threads for efficient computation

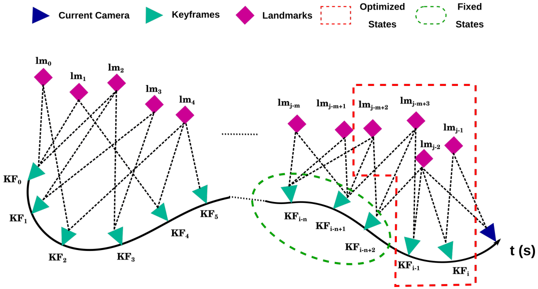


Underwater Monocular Visual SLAM

Local Windowed Bundle Adjustment

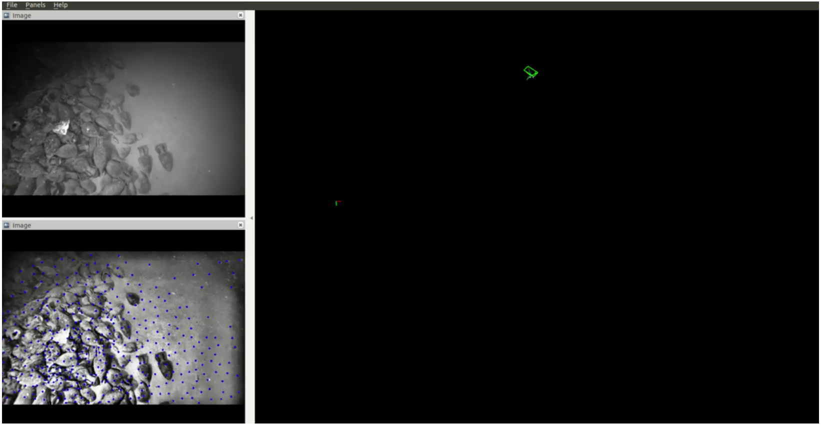
- Optimize most recent keyframes and 3D landmarks by minimization of reprojection errors :

$$\mathcal{X}^* = \underset{\mathcal{X}}{\arg\min} \left(\sum_{i \in \{KF\}} \sum_{j \in \{lm\}} \rho \left(\mathbf{x}_{ij} - \pi \left(\mathbf{X}_i, \mathbf{lm}_j \right) \right) \right) \quad , \quad \mathcal{X} = \left[\mathbf{X}_{KF_i} \quad \mathbf{lm}_j \right]^T$$



Underwater Monocular Visual SLAM

UW-VO for localization during shipwreck exploration



SLAM from a Monocular Vision-based System

Camera

- Monochromatic Camera
- 20 Hz

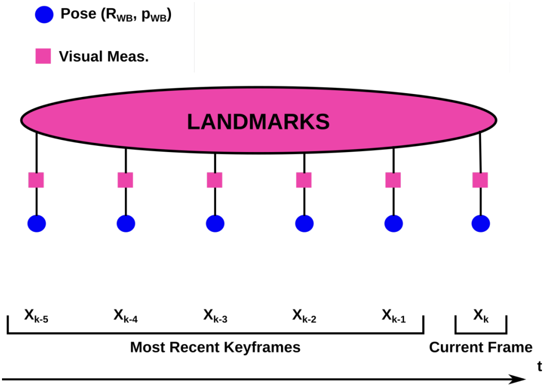
MEMS-IMU

- Inertial meas.
- 200 Hz
- High drift

Pressure Sensor

- Pressure meas. (bar) $\propto$ Depth (m)
- 5-10 Hz

Tight Fusion : Insert other measurement modalities within the factor graph



Tight Fusion : Insert other measurement modalities within the factor graph

Fusion from Maximum Likelihood Estimation

$$\chi^* = \arg\min_{\chi} \left(E_{visual}(\chi) + E_{depth}(\chi) + E_{IMU}(\chi) \right)$$

- E_{visual} : Energy term based on visual measurements
- E_{depth} : Energy term based on pressure measurements
- E_{IMU} : Energy term based on inertial measurements

Visual-Inertial-Pressure SLAM

Low-cost MEMS-IMU Model

- Angular Velocity measurements :

$$\tilde{\boldsymbol{\omega}}_B(t) = \boldsymbol{\omega}_B(t) + \mathbf{b}^g(t) + \boldsymbol{\eta}^g$$

- Linear Acceleration measurements :

$$\tilde{\mathbf{a}}_B(t) = \mathbf{R}_{WB}(t)^T \cdot (\mathbf{a}_W(t) - \mathbf{g}_W) + \mathbf{b}^a(t) + \boldsymbol{\eta}^a$$

Visual-Inertial-Pressure SLAM

Low-cost MEMS-IMU Model

- Angular Velocity measurements :

$$\tilde{\boldsymbol{\omega}}_B(t) = \boldsymbol{\omega}_B(t) + \mathbf{b}^g(t) + \boldsymbol{\eta}^g$$

- Linear Acceleration measurements :

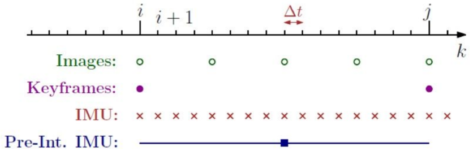
$$\tilde{\mathbf{a}}_B(t) = \mathbf{R}_{WB}(t)^T \cdot (\mathbf{a}_W(t) - \mathbf{g}_W) + \mathbf{b}^a(t) + \boldsymbol{\eta}^a$$

- Measurements corrupted by time-varying biases and zero-mean gaussian noise

Visual-Inertial-Pressure SLAM

IMU Preintegration

- Summarize intra-keyframe IMU measurements as one measurement :



- Relative motion measurements : $\Delta\tilde{\mathbf{R}}_{BiBj}, \Delta\tilde{\mathbf{p}}_{BiBj}, \Delta\tilde{\mathbf{v}}_{BiBj}$

- Easy insertion in the Factor Graph formulation

Visual-Inertial-Pressure SLAM

New state to estimate :

$$\mathbf{X}_i = \begin{bmatrix} \mathbf{R}_{WBi} & \mathbf{p}_{WBi} & \mathbf{v}_{WBi} & \mathbf{b}_i^g & \mathbf{b}_i^a \end{bmatrix}^T$$

IMU Preintegration : Relative errors between keyframes

$$\mathbf{e}_{\Delta\mathbf{R}_{BiBj}} = \hat{\mathbf{R}}_{BiBj} \boxminus \Delta\tilde{\mathbf{R}}_{BiBj}$$
$$\mathbf{e}_{\Delta\mathbf{p}_{BiBj}} = \hat{\mathbf{p}}_{BiBj} - \Delta\tilde{\mathbf{p}}_{BiBj}$$
$$\mathbf{e}_{\Delta\mathbf{v}_{BiBj}} = \hat{\mathbf{v}}_{BiBj} - \Delta\tilde{\mathbf{v}}_{BiBj}$$

$$\mathbf{e}_{\Delta\mathbf{b}^g_{BiBj}} = \hat{\mathbf{b}}^g_{Bj} - \hat{\mathbf{b}}^g_{Bi}$$
$$\mathbf{e}_{\Delta\mathbf{b}^a_{BiBj}} = \hat{\mathbf{b}}^a_{Bj} - \hat{\mathbf{b}}^a_{Bi}$$

Random-walk biases

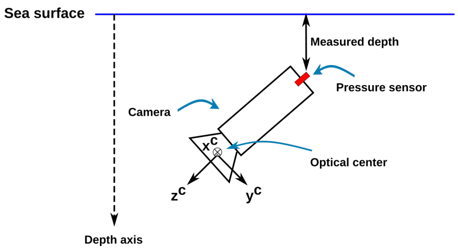
IMU Energy term

$$E_{IMU}(\chi) = \sum_{\mathfrak{K}^*} \left(\mathbf{e}_{imu}(\mathbf{X}_i, \mathbf{X}_j)^T \cdot {\mathbf{\Sigma}^{imu}_{BiBj}}^{-1} \cdot \mathbf{e}_{imu}(\mathbf{X}_i, \mathbf{X}_j) \right)$$

$$\mathbf{e}_{imu}(\mathbf{X}_i, \mathbf{X}_j) = \begin{bmatrix} \mathbf{e}_{\Delta\mathbf{R}_{BiBj}} & \mathbf{e}_{\Delta\mathbf{p}_{BiBj}} & \mathbf{e}_{\Delta\mathbf{v}_{BiBj}} & \mathbf{e}_{\Delta\mathbf{b}^g_{BiBj}} & \mathbf{e}_{\Delta\mathbf{b}^a_{BiBj}} \end{bmatrix}^T$$

Visual-Inertial-Pressure SLAM

Pressure Sensor

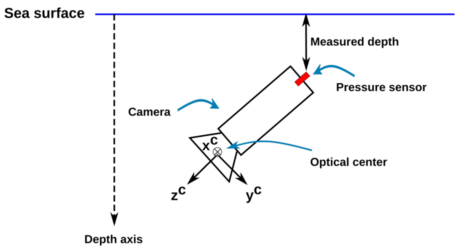


- Pressure measurements : pressure (Pa) $\propto$ depth (m)
- Depth variation from starting point :

$$\tilde{d}_i = {}_{\text{raw}}\tilde{d}_i - {}_{\text{raw}}\tilde{d}_0$$

Visual-Inertial-Pressure SLAM

Pressure Sensor



- Pressure measurements : pressure (Pa) $\propto$ depth (m)
- Depth variation from starting point :

$$\tilde{d}_i = {}_{\text{raw}}\tilde{d}_i - {}_{\text{raw}}\tilde{d}_0$$

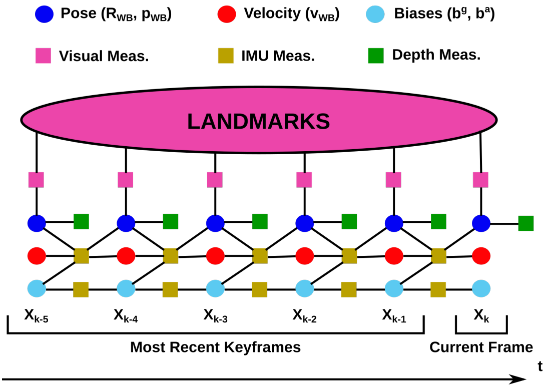
Integration of absolute depth measurements :

$$E_{depth}(\mathbf{X}_i) = \|\tilde{d}_i - \hat{t}^z{}_{wc_i}\|^2_{\sigma^2_{depth}}$$

Visual-Inertial-Pressure SLAM

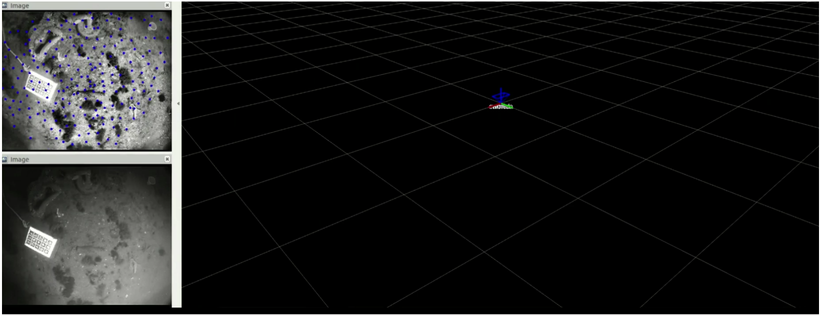
Visual-Inertial-Pressure Optimization

$$\chi^* = \arg\min_{\chi} \left(E_{visual}(\chi) + E_{IMU}(\chi) + E_{depth}(\chi) \right)$$



Visual-Inertial-Pressure SLAM

UW-VIP for localization with short loss of visual information



Monocular Dense 3D Mapping

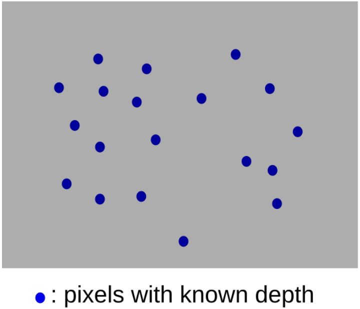
Dense 3D Mapping

- Densify the sparse 3D measurements
- Make use of optimized states : keyframes + 3D landmarks

Monocular Dense 3D Mapping

Depth Map Densification

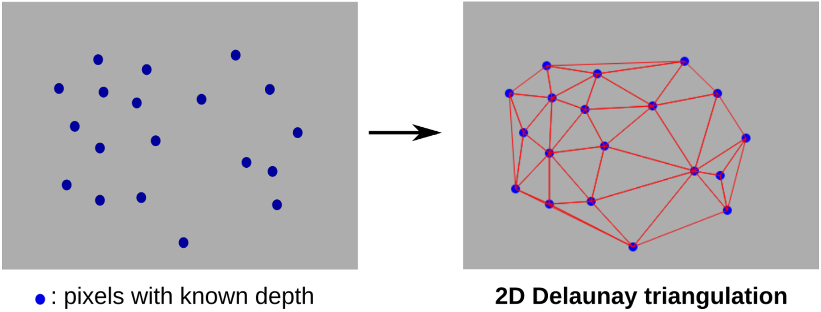
- Find 3D features nearest-neighbors from 2D Delaunay triangulation



Monocular Dense 3D Mapping

Depth Map Densification

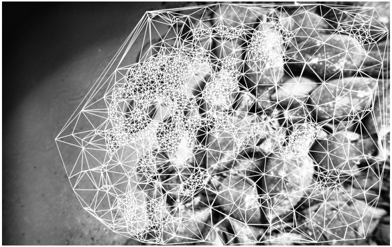
- Find 3D features nearest-neighbors from 2D Delaunay triangulation
- Depth value interpolation from Delaunay triangles

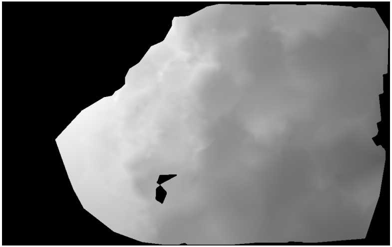


Monocular Dense 3D Mapping

Depth Map Densification

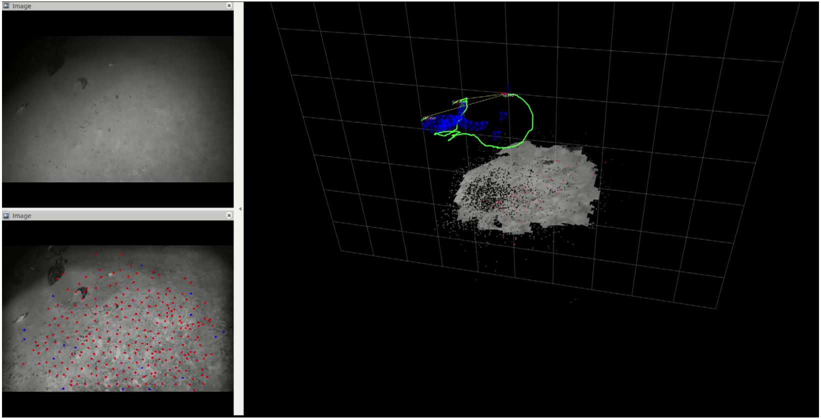
- Find 3D features nearest-neighbors from 2D Delaunay triangulation
- Depth value interpolation from Delaunay triangles


(a) 2D Delaunay triangulation.


(b) 2D densified depth map.

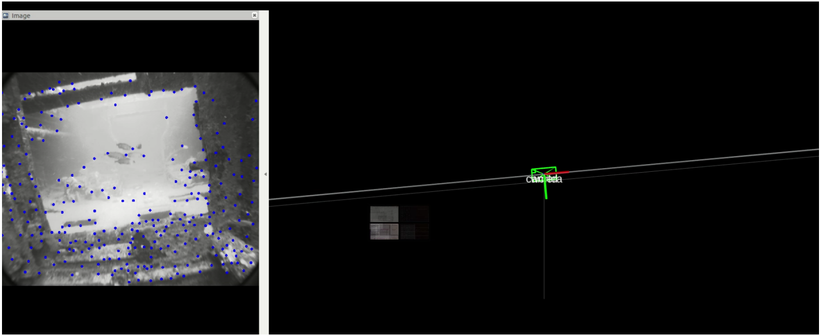
Monocular Dense 3D Mapping

Online 3D Reconstruction



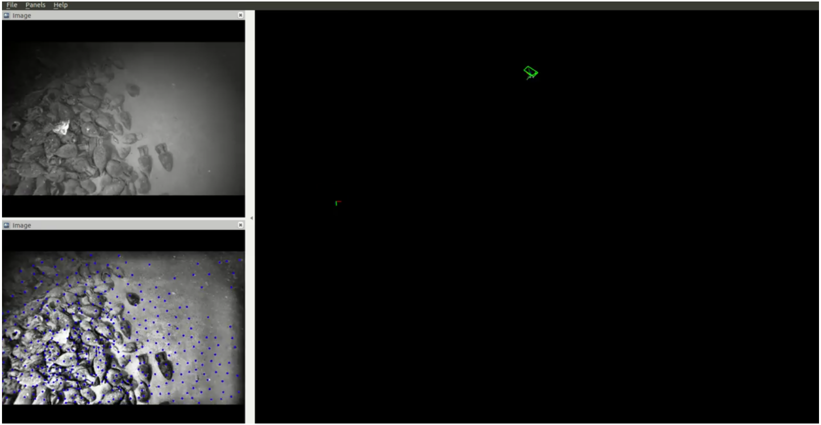
Monocular Dense 3D Mapping

Online 3D Reconstruction in Complex Environment



Monocular Dense 3D Mapping

Post-mission 3D Reconstruction

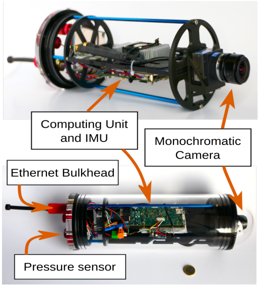


Conclusion

Conclusion

Experimental Validation

- Algorithms validated on the Tegra TX2
- All the methods run in real-time
- Release of a public dataset : AQUALOC



Conclusion

AQUALOC Dataset : `http://www.lirmm.fr/aqualoc/`

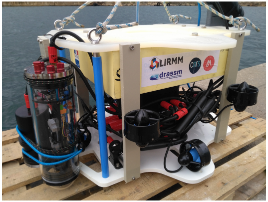


FIGURE – ROV Dumbo (DRASSM / LIRMM)

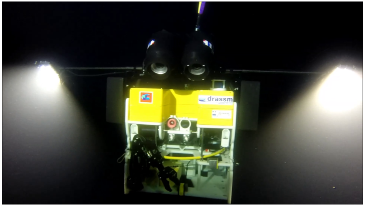
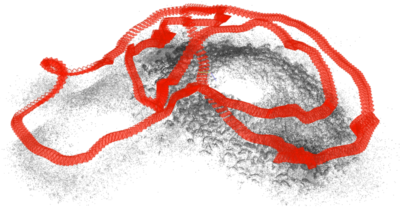


FIGURE – ROV Perseo (Copetech SM - Credit : DRASSM / F. Osada)

- 17 sequences
- Synchronized measurements
- Harbor & Archaeological sites
- Comparative baselines from offline photogrammetry



Conclusion

Follow-up

- Collaboration with Stanford Robotics Lab on OceanOne

What's next ?

- Add a loop-closure feature
- Manage localization updates in the 3D reconstruction

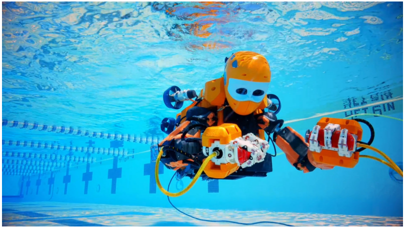


FIGURE – Ocean One - Stanford University

Related Publications I

Marcela Carvalho, Maxime Ferrera, Alexandre Boulch, Julien Moras, Bertrand Le Saux et Pauline Trouvé-Peloux. "Technical Report : Co-learning of geometry and semantics for online 3D mapping". In : arXiv preprint arXiv :1911.01082 (2019).

Maxime Ferrera, Alexandre Boulch et Julien Moras. "Fast Stereo Disparity Maps Refinement By Fusion of Data-Based And Model-Based Estimations". In : 3DV. 2019.

Maxime Ferrera, Vincent Creuze, Julien Moras et Pauline Trouvé-Peloux. "AQUALOC : An Underwater Dataset for Visual-Inertial-Pressure Localization.". In : The International Journal of Robotics Research. 2019.

Maxime Ferrera, Julien Moras, Pauline Trouvé-Peloux et Vincent Creuze. "Localisation autonome basée vision d'un robot sous-marin et cartographie de précision". In : ORASIS. 2017.

Related Publications II

Maxime FERRERA, Julien MORAS, Pauline TROUVÉ-PELOUX et Vincent CREUZE. "Odométrie Visuelle Monoculaire en Environnement Sous-Marin". In : Reconnaissance des Formes, Image, Apprentissage et Perception (RFIAP). 2018.

Maxime FERRERA, Julien MORAS, Pauline TROUVÉ-PELOUX et Vincent CREUZE. "Real-Time Monocular Visual Odometry for Turbid and Dynamic Underwater Environments". In : Sensors. T. 19. 3. 2019.

Maxime FERRERA, Julien MORAS, Pauline TROUVÉ-PELOUX, Vincent CREUZE et Denis DÉGEZ. "The Aqualoc Dataset: Towards Real-Time Underwater Localization from a Visual-Inertial-Pressure Acquisition System". In : IROS Workshop - New Horizons for Underwater Intervention Missions : from Current Technologies to Future Applications. 2018.

THANK YOU!

Contact : `maxime.ferrera@gmail.com`
Website : `https://ferreram.github.io/`
More videos : *Aqualoc channel on Youtube*