

Odométrie Visuelle Monoculaire en Environnement Sous-Marin

Maxime Ferrera^{1,2}

Doctorant

Encadrants : Julien Moras¹

Pauline Trouvé-Peloux¹

Directeur : Vincent Creuze²

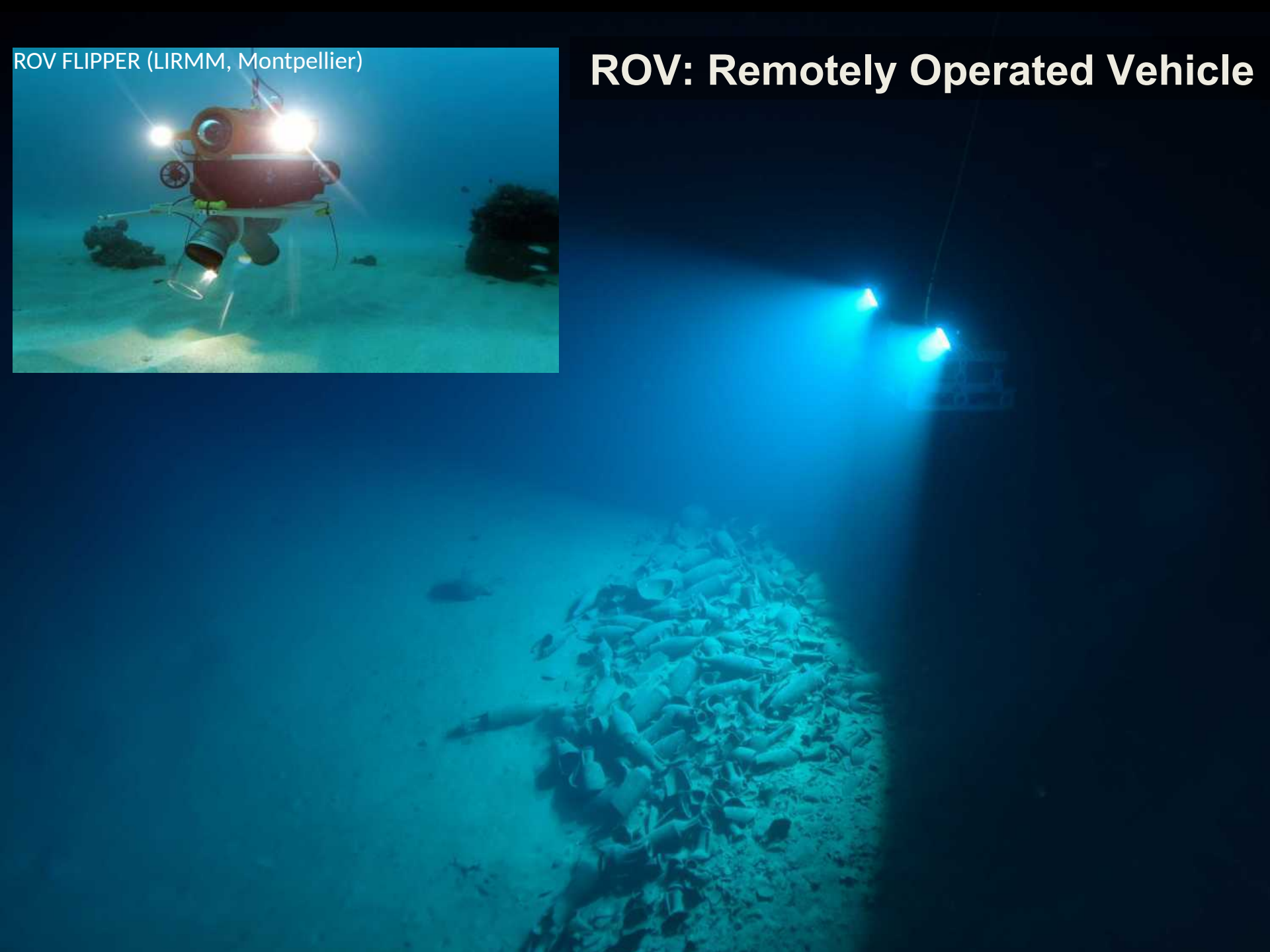
¹ONERA, The French Aerospace Lab

²LIRMM, Université de Montpellier, CNRS

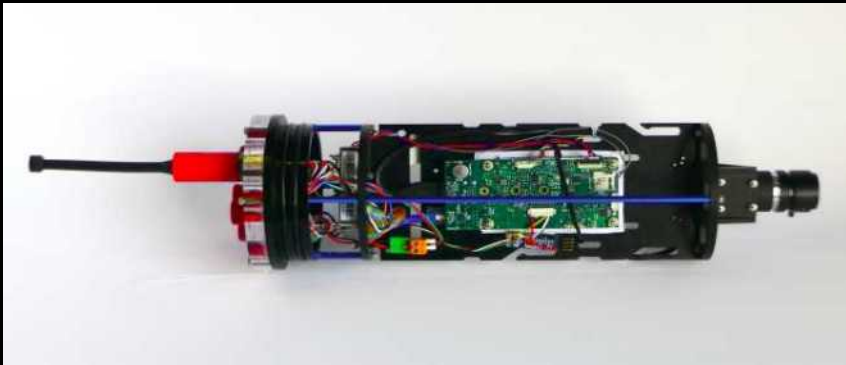
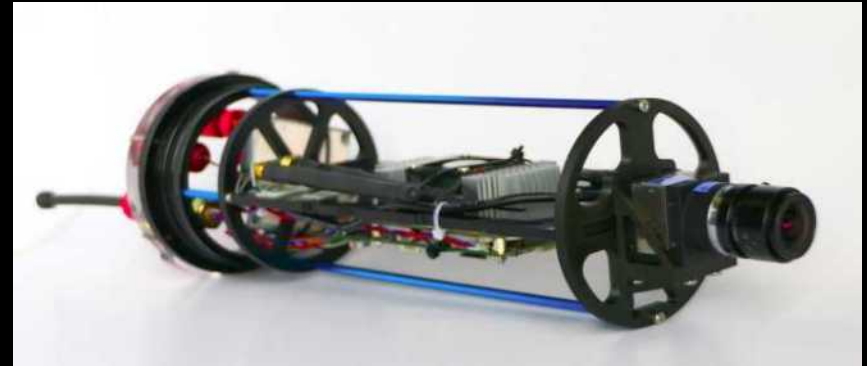
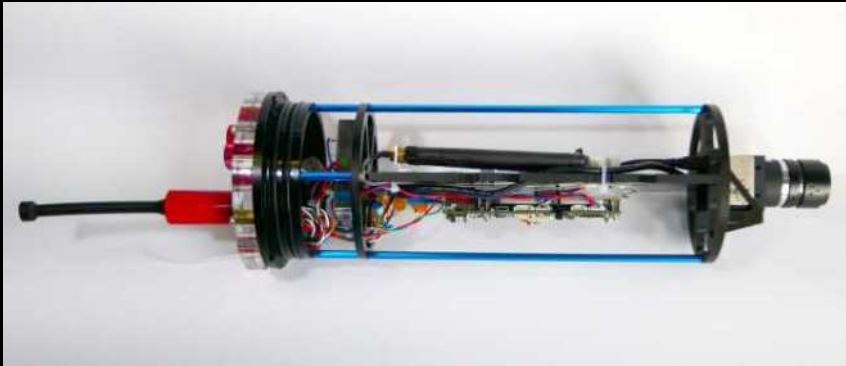
ROV FLIPPER (LIRMM, Montpellier)



ROV: Remotely Operated Vehicle



Charge Utile



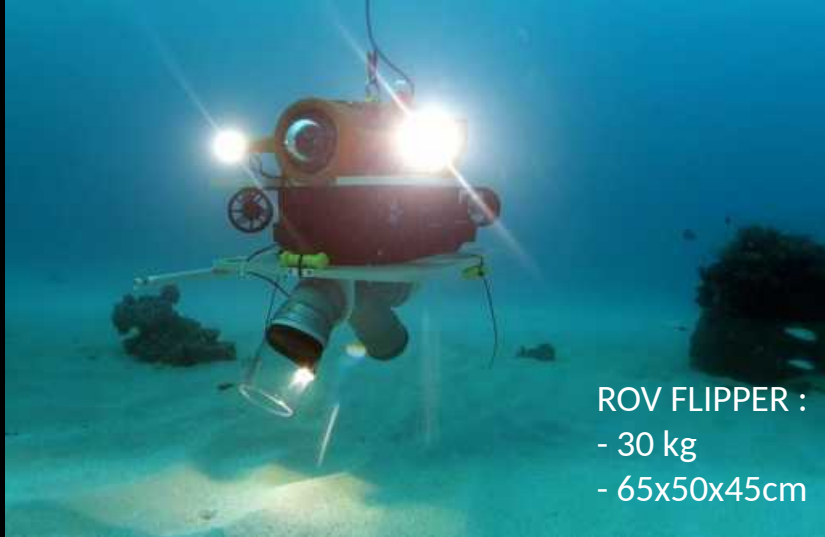
Caractéristiques :

- Dimensions : 33.4 x 11.4 cm
- Hublot : dôme sphérique

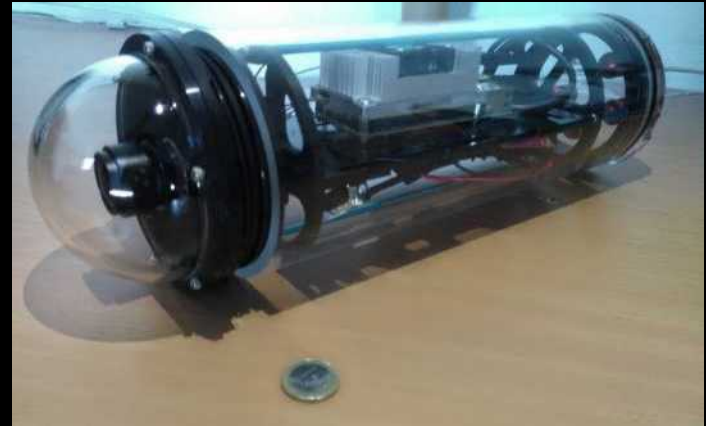
Instrumentation :

- Caméra monochrome
- IMU – MEMS
- Capteur de pression
- Calculateur Tegra TX2

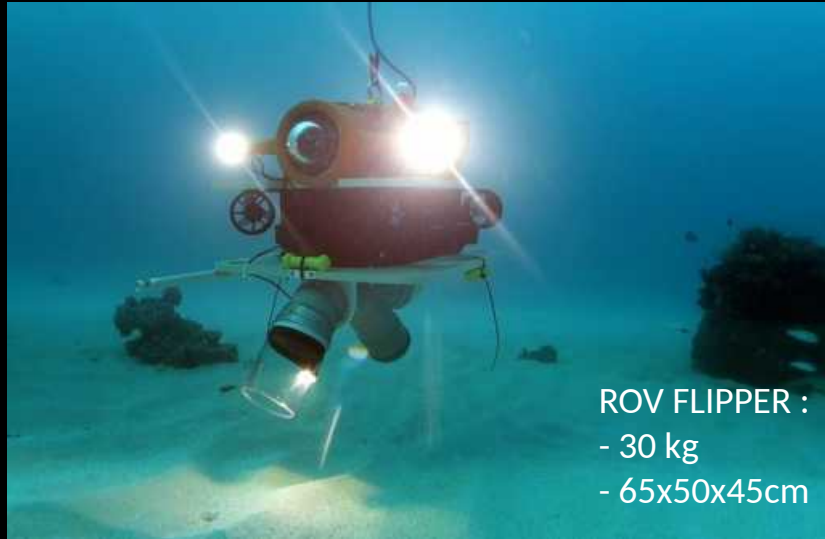
Objectif



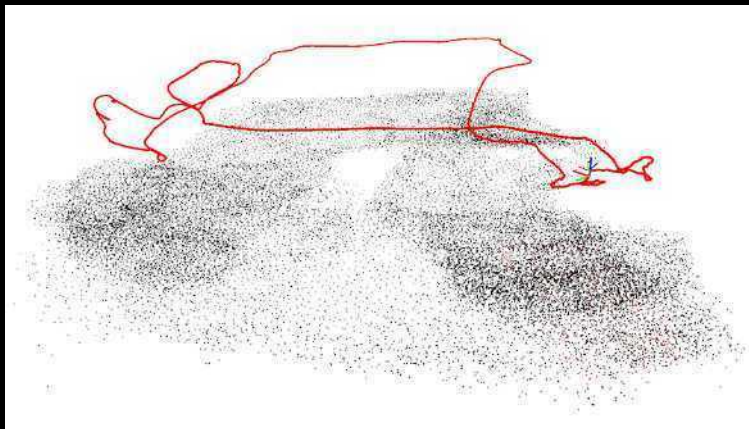
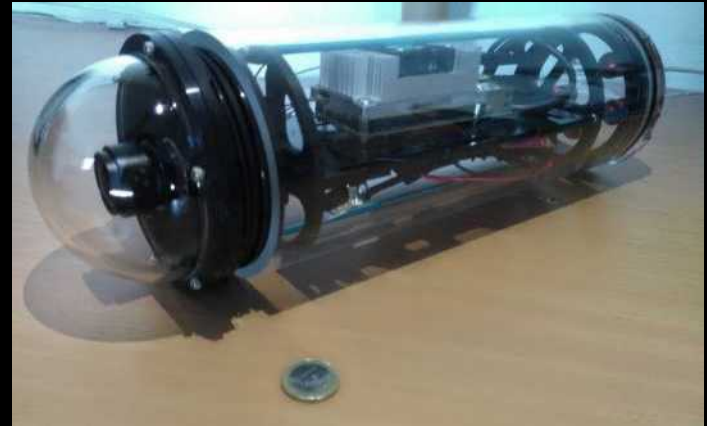
Caméra + Calculateur Embarqué



Objectif



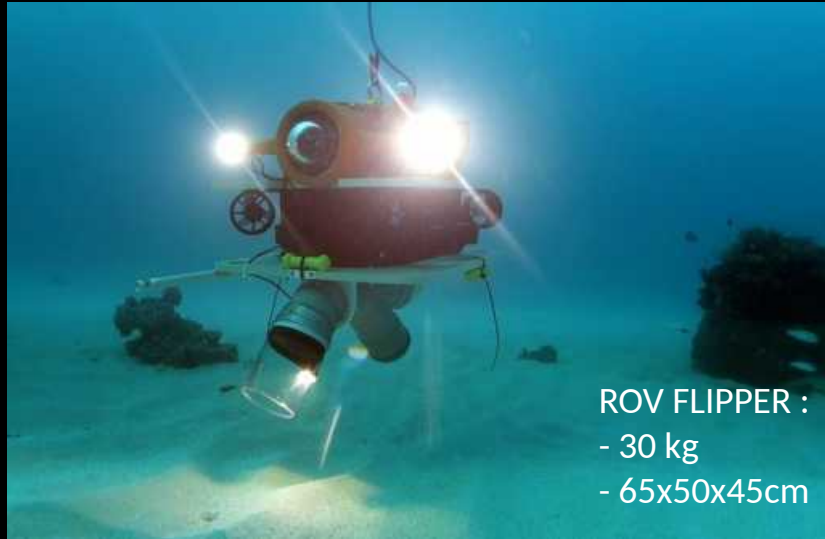
Caméra + Calculateur Embarqué



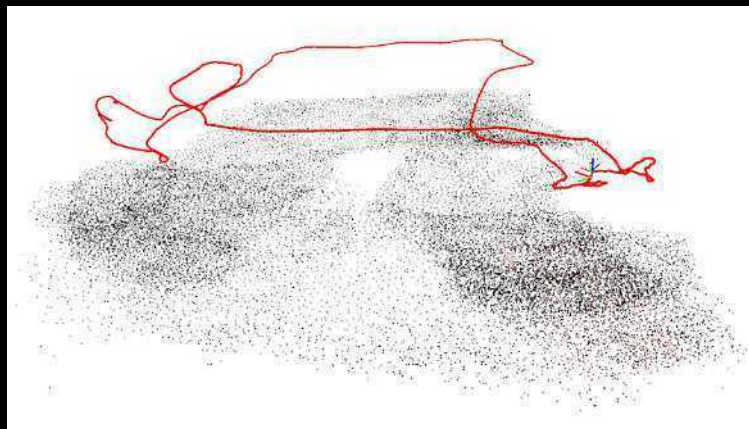
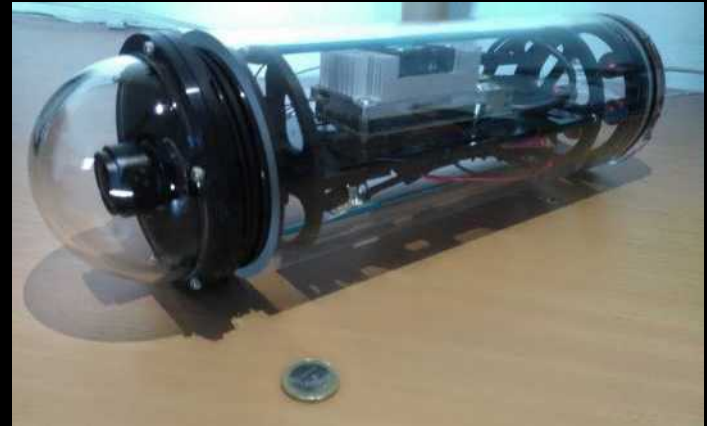
→ **Trajectoire**

→ **Carte 3D**

Objectif



Caméra + Calculateur Embarqué



Trajectoire



Carte 3D



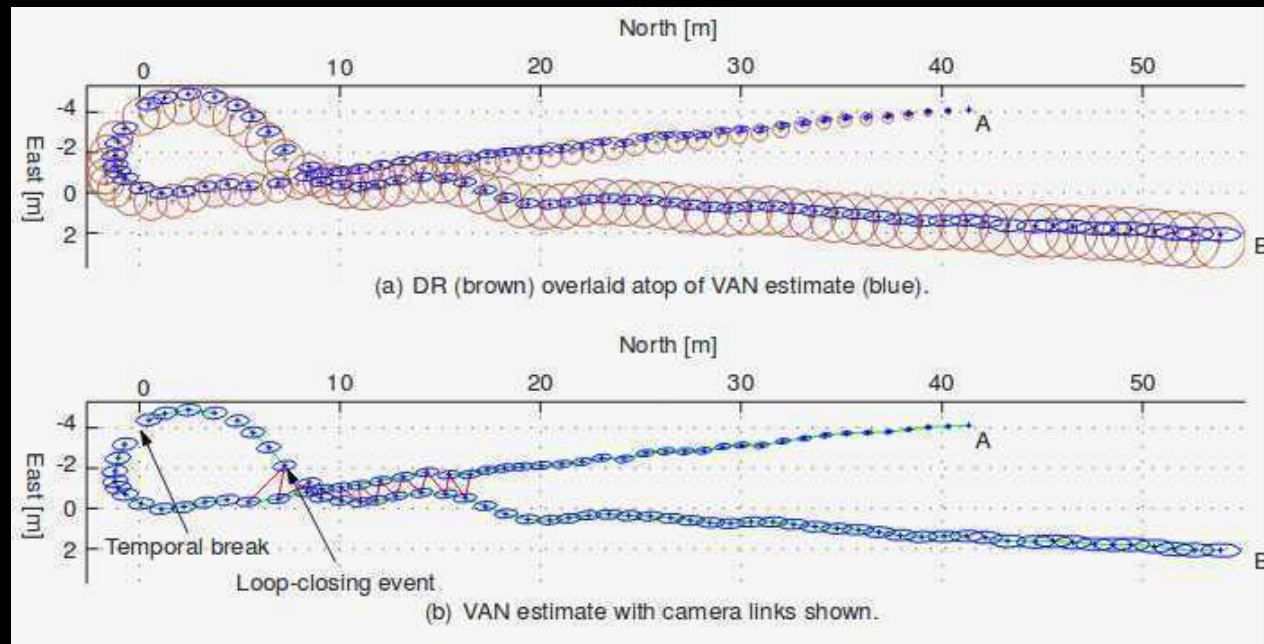
Odométrie Visuelle

—

SLAM visuel

État de l'art

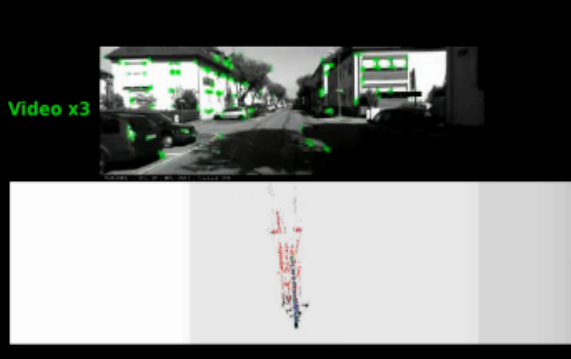
- Robotique sous-marine :
 - Méthodes de filtrages (EKF, EIF) :
 - Fusion de capteurs de navigation
 - Utilisation de caméras pour limiter la dérive



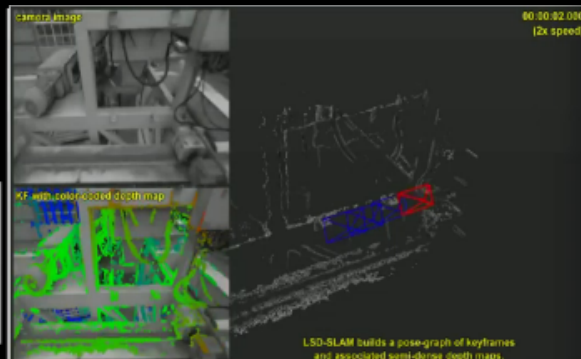
État de l'art

- Robotique terrestre et aérienne :
 - Fort intérêt pour les méthodes de localisation par vision monoculaire
 - Nombreux algorithmes open-source :

ORB-SLAM¹



LSD-SLAM²



SVO³



1: R. Mur-Artal, J. M. M. Montiel, and J. D. Tardós. ORB-SLAM: A Versatile and Accurate Monocular SLAM System. *IEEE Transactions on Robotics*, 2015.

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

3: Forster, C., Pizzoli, M., Scaramuzza, D., 2017. SVO: Fast semi-direct monocular visual odometry. *ICRA 2014*.

Vision sous-marine v.s. Vision terrestre



Séquence Urbaine
(KITTI¹)



Séquence sous-marine
(crédit : DRASSM²)

1: A. Geiger, P. Lenz, and R. Urtasun. Are we ready for Autonomous Driving? The KITTI Vision Benchmark Suite. In CVPR, 2012.

2 : Département de Recherche en Archéologie Subaquatique et Sous-Marine - Ministère de la Culture

Vision sous-marine v.s. Vision terrestre

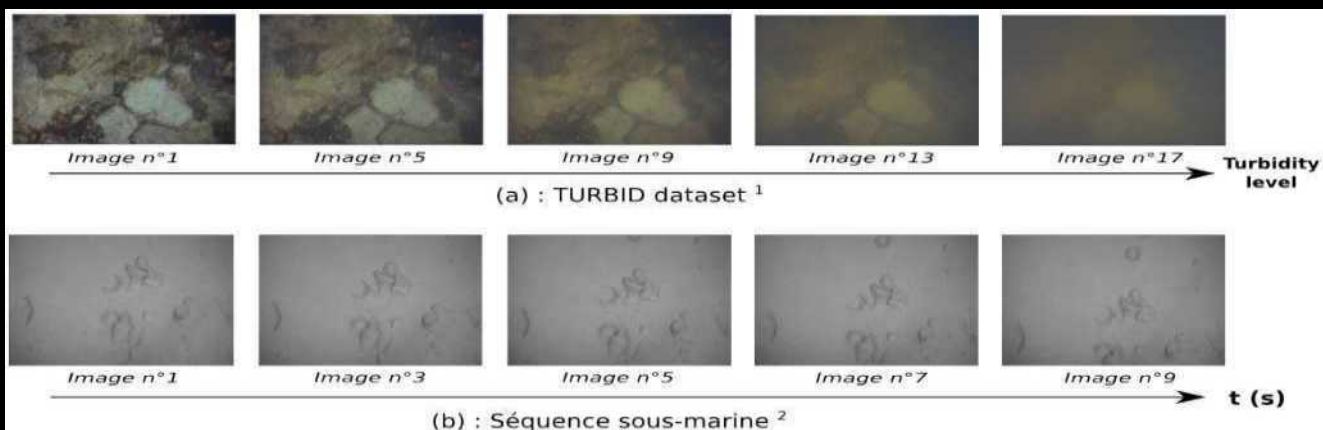
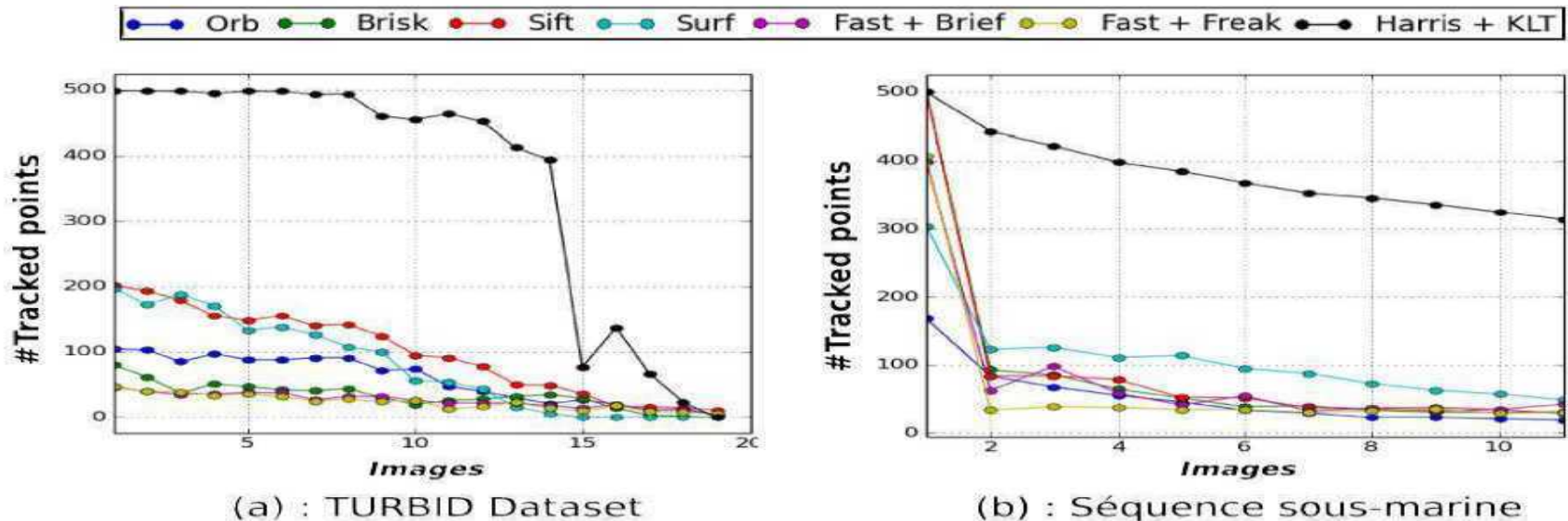


- ✗ Texture répétitive
- ✗ Turbidité
- ✗ Environnement dynamique
- ✗ Éclairage artificiel

Contributions

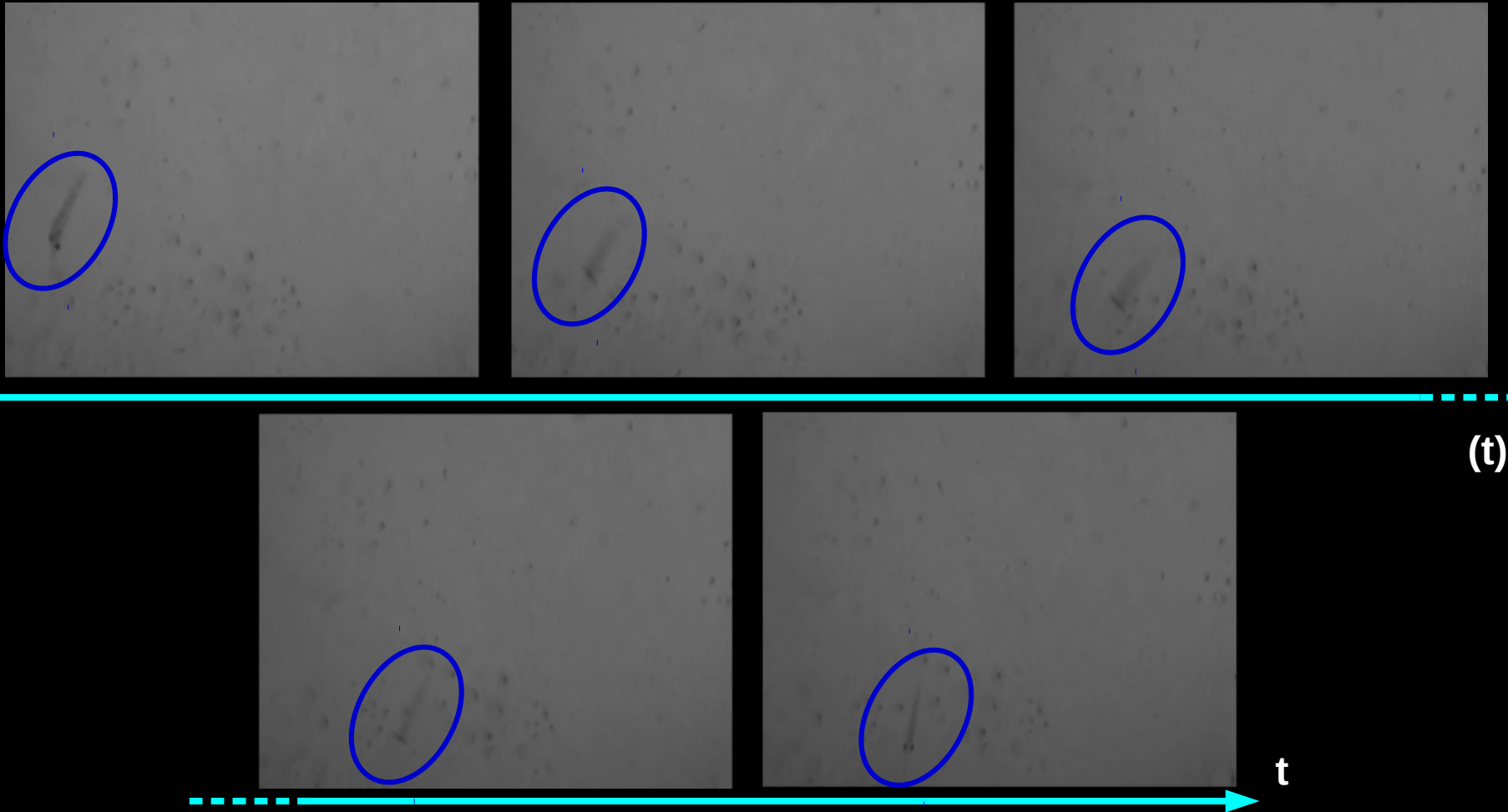
- Évaluation des méthodes de suivi de points d'intérêts sur des données sous-marines
- Développement de UW-VO :
 - algorithme d'odométrie monoculaire basée sur un suivi par flot optique
- Développement d'un mécanisme de recherche des points perdus par flot optique
- Comparaison avec l'état de l'art en SLAM et odométrie visuelle

Évaluation des méthodes de suivi de points



- ✓ **Suivi image à image par KLT**

Mécanisme de recherche des points perdus



- Environnement dynamique

⇒ occultations fréquentes¹⁵!

Mécanisme de recherche des points perdus

Paramètres d'acquisition

- Résolution des images : 640x480 pixels
- Fréquence d'acquisition des images : 20 hz
- Distance de travail : 0 à 3m
- Surface imagée : $\sim 3 \times 2$ m (à 3m)

Caractérisation des dynamiques

- Vitesse de déplacement du robot v_{rob} : 0 à 2m/s ($\sim 7\text{km/h}$)
- Vitesse de déplacement des objets dynamiques : $v_{\text{dyn}} \gg v_{\text{rob}}$

Mécanisme de recherche des points perdus

Paramètres d'acquisition

- Résolution des images : 640x480 pixels
- Fréquence d'acquisition des images : 20 hz
- Distance de travail : 0 à 3m
- Surface imagée : $\sim 3 \times 2$ m (à 3m)

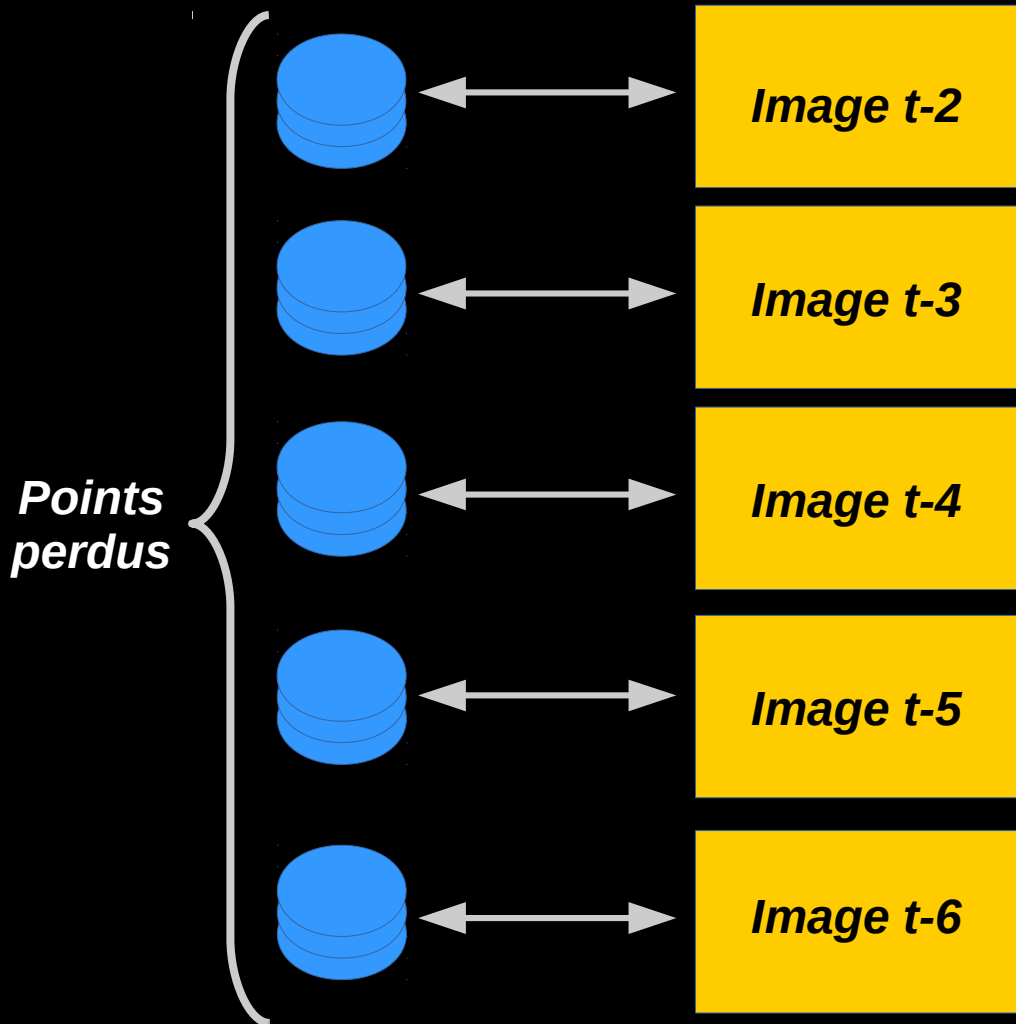
Caractérisation des dynamiques

- Vitesse de déplacement du robot v_{rob} : 0 à 2m/s (~ 7 km/h)
- Vitesse de déplacement des objets dynamiques : $v_{\text{dyn}} \gg v_{\text{rob}}$

- Déplacement max par image : 0.1 m $\Leftrightarrow$ ~ 20 px (à 3m)
- Déplacement des objets dynamiques : > 20 px / image
- Fenêtre du KLT : 21×21 px

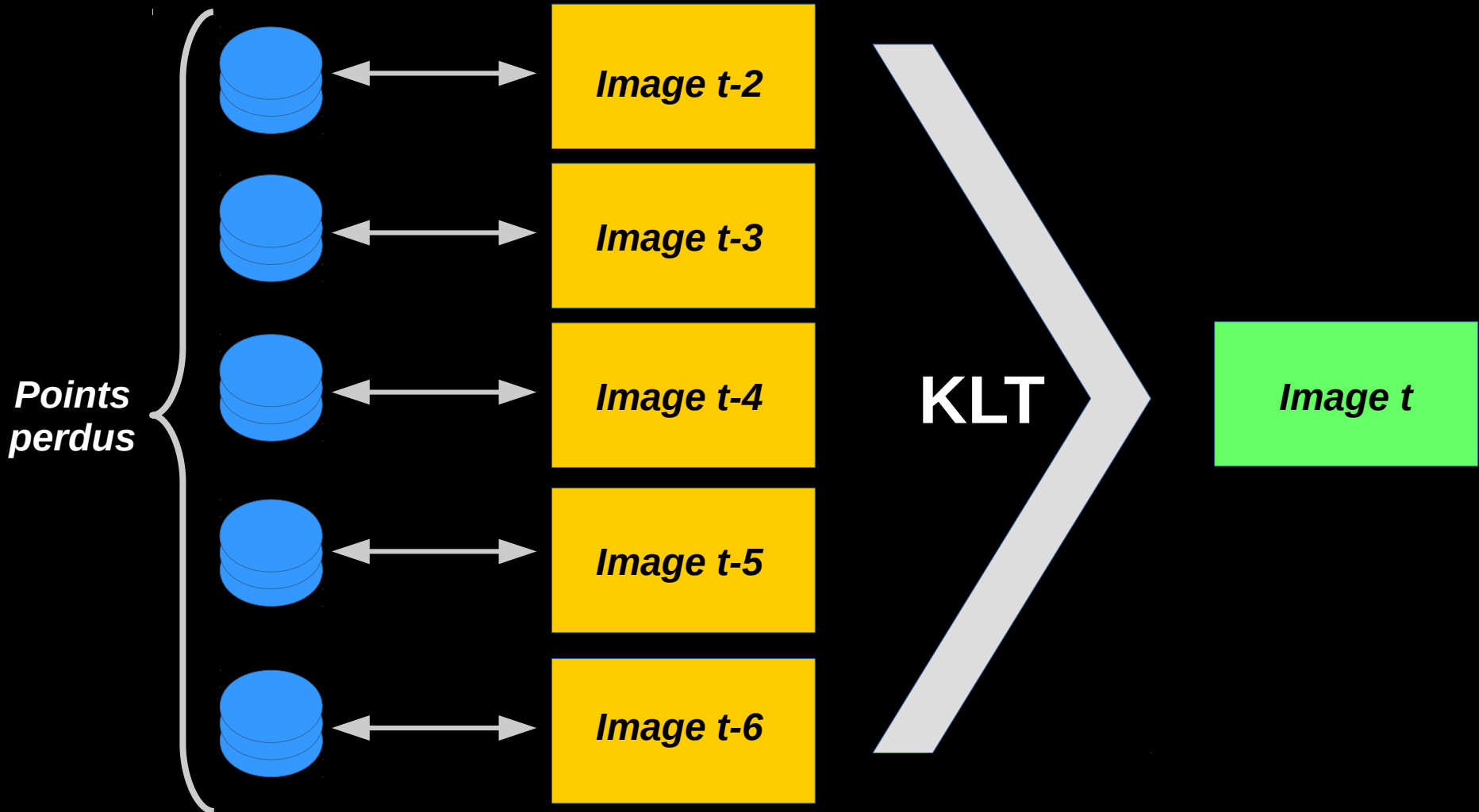
$\Rightarrow$ ***Bassin de convergence du KLT étendu sur plusieurs images***

Mécanisme de recherche des points perdus

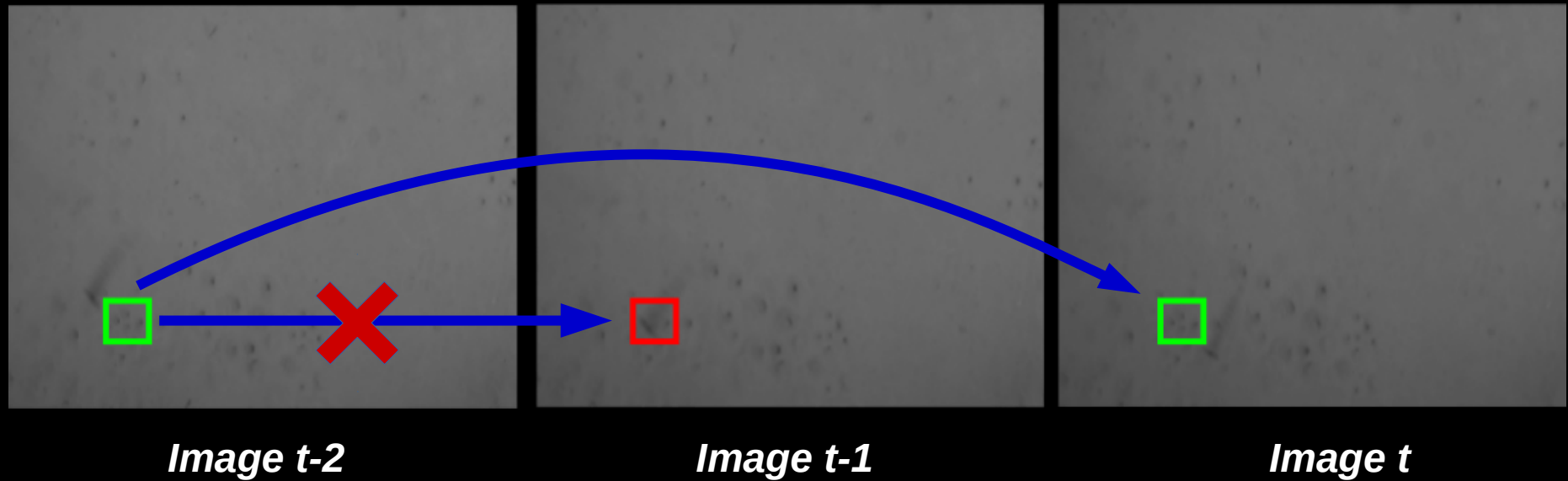


- Mise en mémoire des points perdus
- Recherche des points sur une fenêtre glissante d'images

Mécanisme de recherche des points perdus

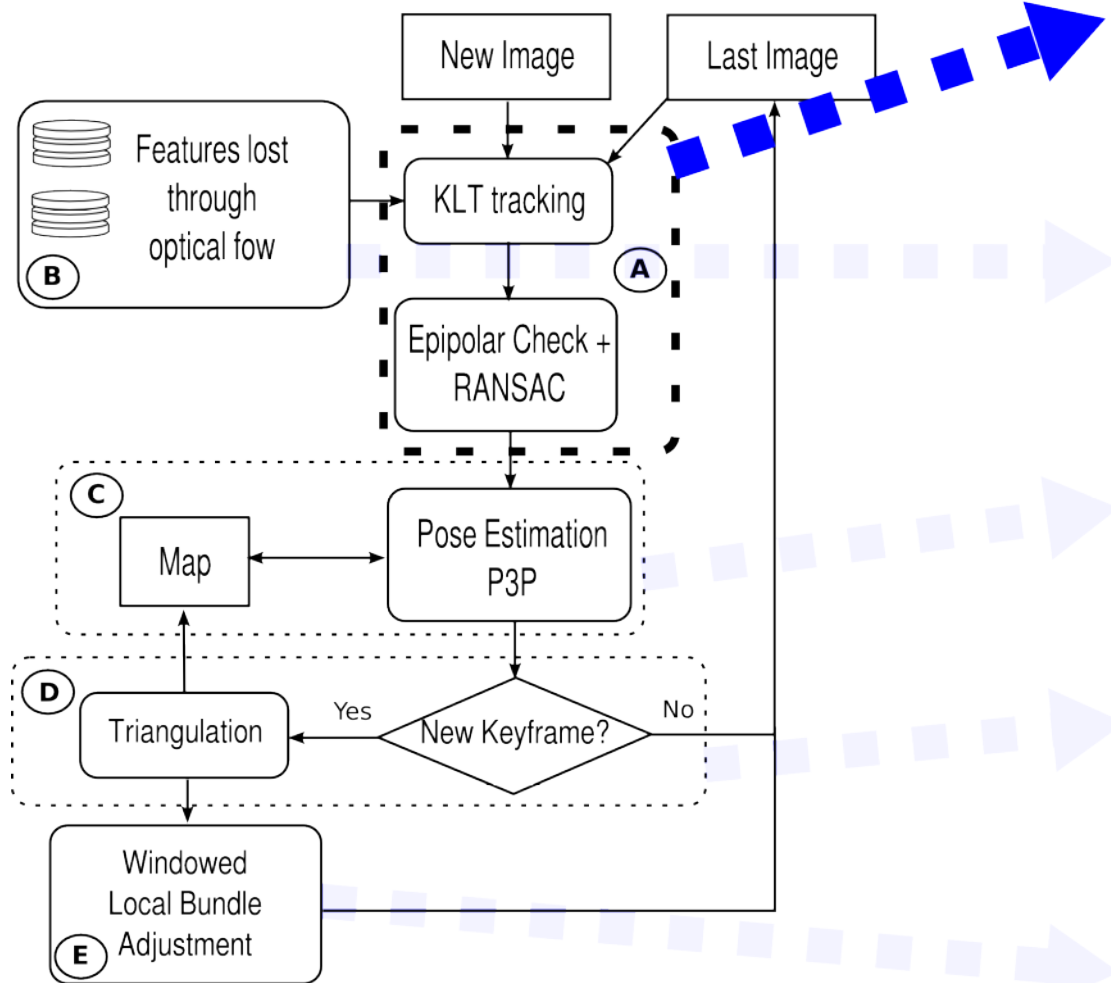


Mécanisme de recherche des points perdus



- ✓ Suivi image à image par KLT
- ✓ **Recherche multi-vues des points occultés par KLT sur fenêtre glissante**

Pipeline de UW-VO



(A)

Méthode basée flot optique :

- + : Efficace pour le suivi de points peu discriminants
- : Suivi image à image seulement

(B)

Recherche des points perdus :

- Mise en mémoire des points récemment perdus
- Fenêtre glissante des cinq dernières images
- + : Robustesse aux occultations temporaires

(C)

Estimation de la pose 6D (t + R) :

- Calcul de la pose à partir du triplet de correspondances 3D-2D le plus vraisemblable

(D)

SLAM basé images-clés :

- Estimation des points 3D quand assez de disparité
- Enregistrement des images-clés et des points 3D observés

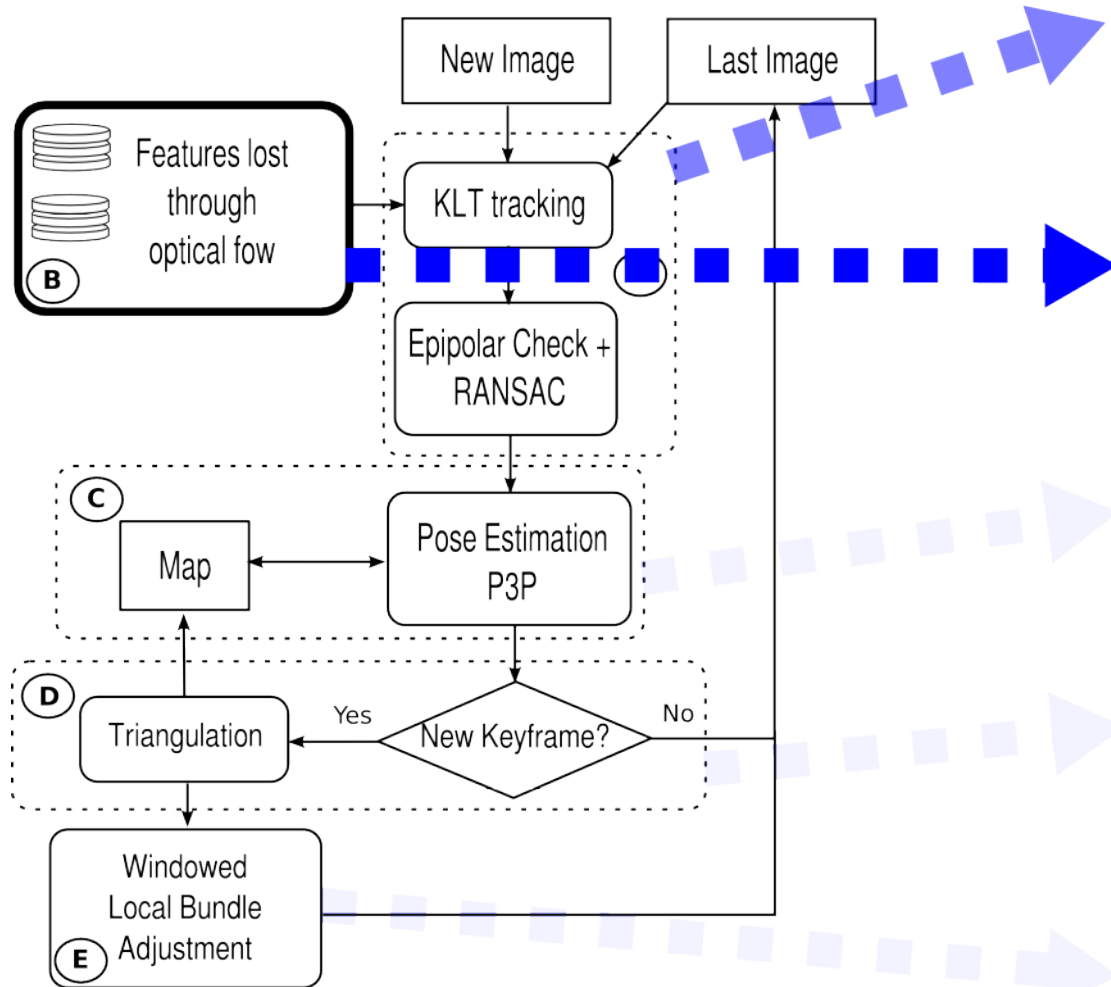
(E)

Optimisation de la trajectoire :

- Minimisation de l'erreur de reprojection sur une fenêtre d'images-clés (Bundle Adjustment) :

$$\arg \min_{P_i, X_j} \left(\sum_i \sum_j (x_{ij} - \pi_i(X_j))^2 \right)$$

Pipeline de UW-VO



(A)

Méthode basée flot optique :

- + : Efficace pour le suivi de points peu discriminants
- : Suivi image à image seulement

(B)

Recherche des points perdus :

- Mise en mémoire des points récemment perdus
- Fenêtre glissante des cinq dernières images
- + : Robustesse aux occultations temporaires

(C)

Estimation de la pose 6D (t + R) :

- Calcul de la pose à partir du triplet de correspondances 3D-2D le plus vraisemblable

(D)

SLAM basé images-clés :

- Estimation des points 3D quand assez de disparité
- Enregistrement des images-clés et des points 3D observés

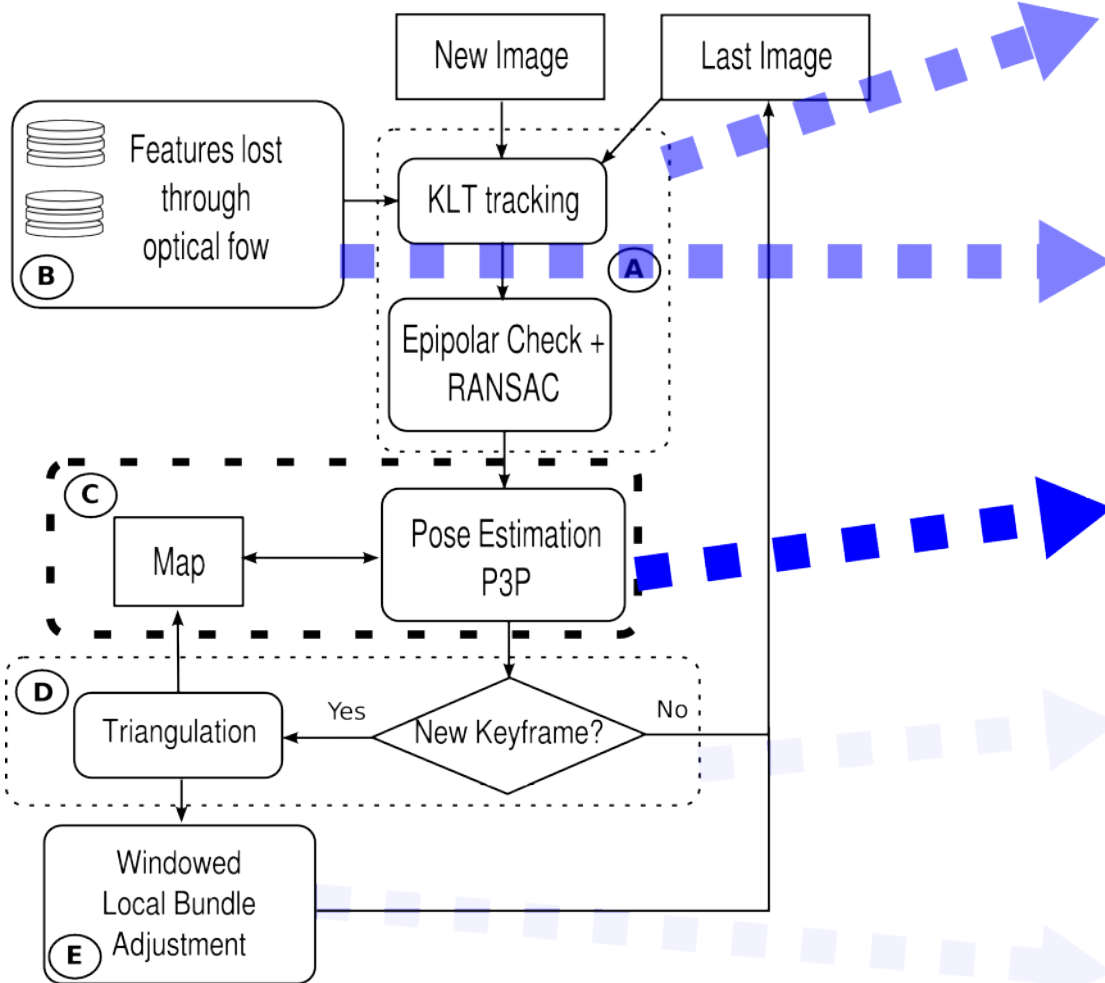
(E)

Optimisation de la trajectoire :

- Minimisation de l'erreur de reprojection sur une fenêtre d'images-clés (Bundle Adjustment) :

$$\arg \min_{P_i, X_j} \left(\sum_i \sum_j (x_{ij} - \pi_i(X_j))^2 \right)$$

Pipeline de UW-VO



(A)

Méthode basée flot optique :

- + : Efficace pour le suivi de points peu discriminants
- : Suivi image à image seulement

(B)

Recherche des points perdus :

- Mise en mémoire des points récemment perdus
- Fenêtre glissante des cinq dernières images
- + : Robustesse aux occultations temporaires

(C)

Estimation de la pose 6D (t + R) :

- Calcul de la pose à partir du triplet de correspondances 3D-2D le plus vraisemblable

(D)

SLAM basé images-clés :

- Estimation des points 3D quand assez de disparité
- Enregistrement des images-clés et des points 3D observés

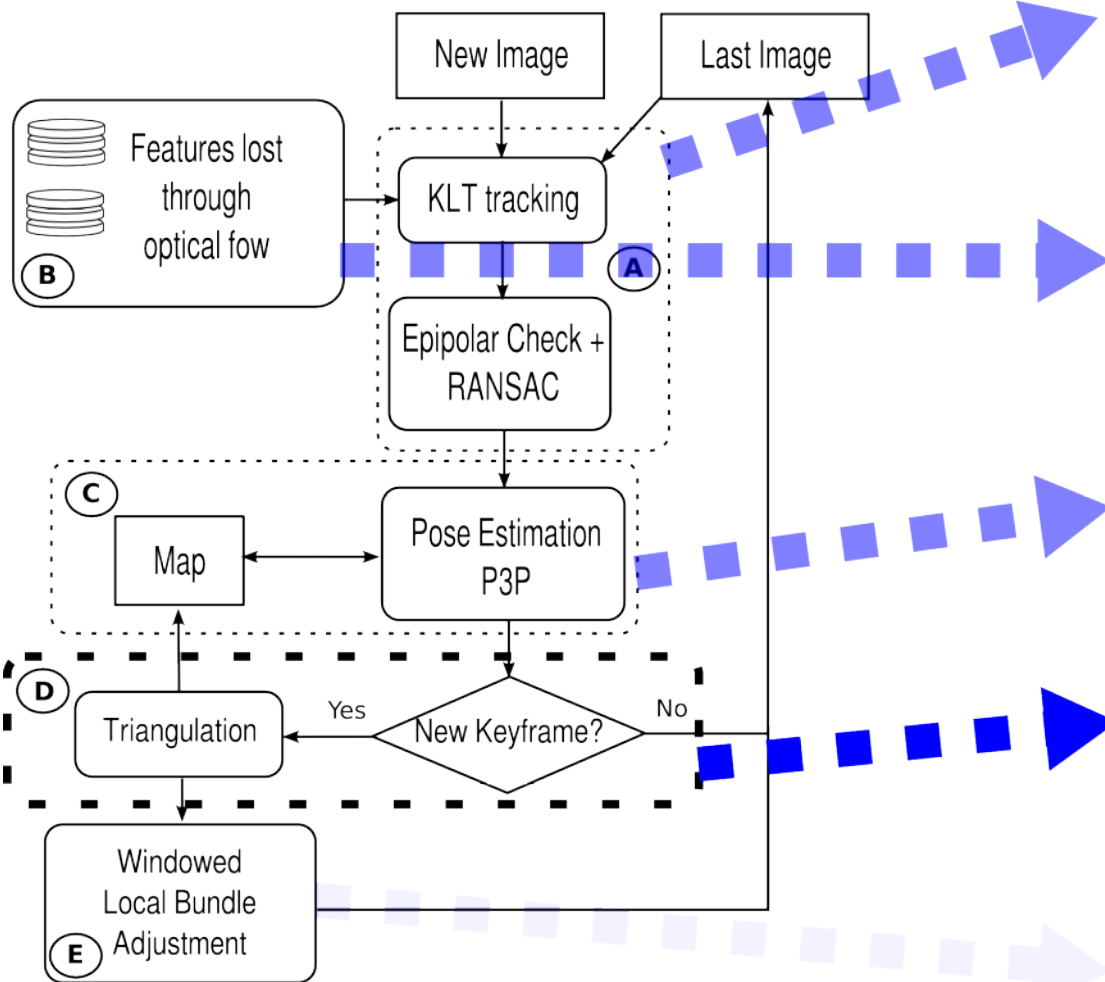
(E)

Optimisation de la trajectoire :

- Minimisation de l'erreur de reprojection sur une fenêtre d'images-clés (Bundle Adjustment) :

$$\arg \min_{P_i, X_j} \left(\sum_i \sum_j (x_{ij} - \pi_i(X_j))^2 \right)$$

Pipeline de UW-VO



(A)

Méthode basée flot optique :

- + : Efficace pour le suivi de points peu discriminants
- : Suivi image à image seulement

(B)

Recherche des points perdus :

- Mise en mémoire des points récemment perdus
- Fenêtre glissante des cinq dernières images
- + : Robustesse aux occultations temporaires

(C)

Estimation de la pose 6D (t + R) :

- Calcul de la pose à partir du triplet de correspondances 3D-2D le plus vraisemblable

(D)

SLAM basé images-clés :

- Estimation des points 3D quand assez de disparité
- Enregistrement des images-clés et des points 3D observés

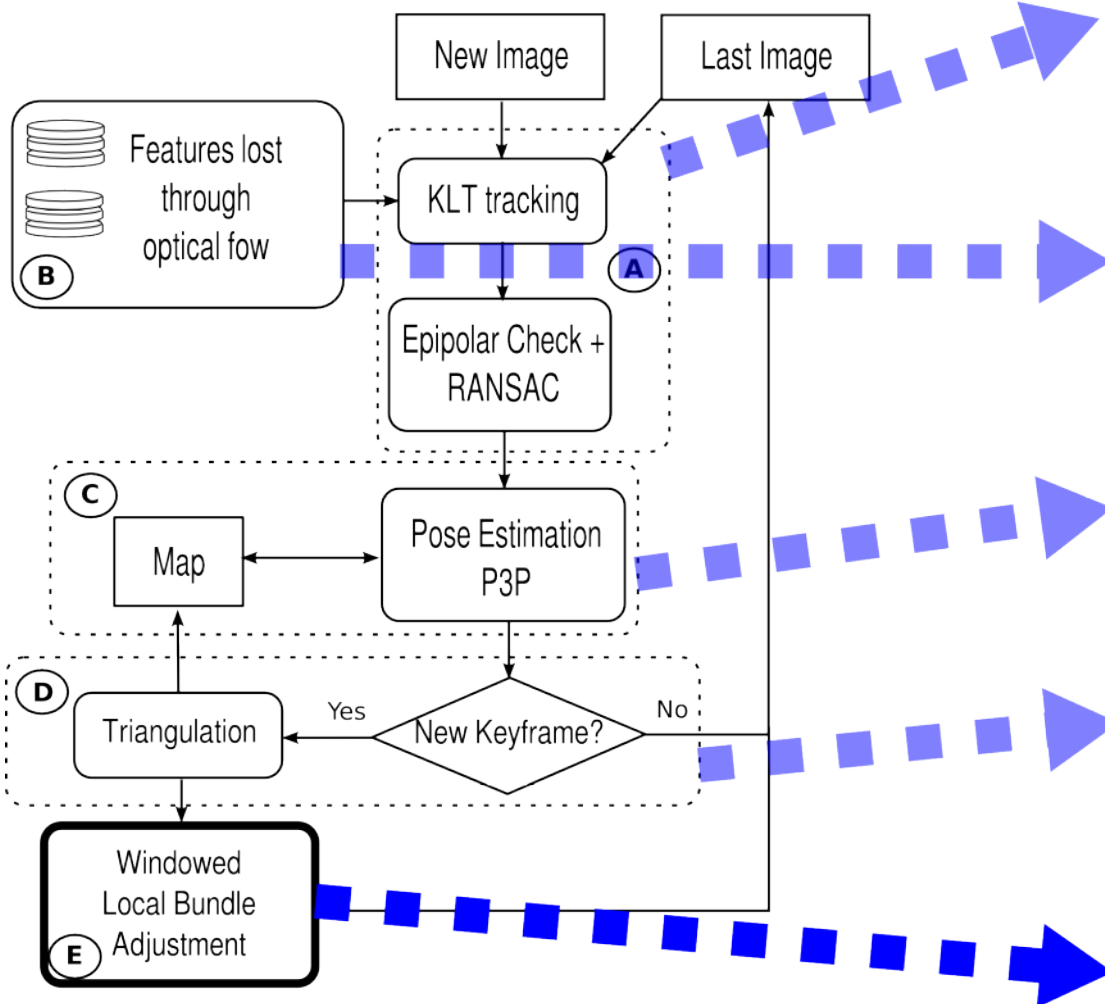
(E)

Optimisation de la trajectoire :

- Minimisation de l'erreur de reprojection sur une fenêtre d'images-clés (Bundle Adjustment) :

$$\arg \min_{P_i, X_j} \left(\sum_i \sum_j (x_{ij} - \pi_i(X_j))^2 \right)$$

Pipeline de UW-VO



(A)

Méthode basée flot optique :

- + : Efficace pour le suivi de points peu discriminants
- : Suivi image à image seulement

(B)

Recherche des points perdus :

- Mise en mémoire des points récemment perdus
- Fenêtre glissante des cinq dernières images
- + : Robustesse aux occultations temporaires

(C)

Estimation de la pose 6D (t + R) :

- Calcul de la pose à partir du triplet de correspondances 3D-2D le plus vraisemblable

(D)

SLAM basé images-clés :

- Estimation des points 3D quand assez de disparité
- Enregistrement des images-clés et des points 3D observés

(E)

Optimisation de la trajectoire :

- Minimisation de l'erreur de reprojection sur une fenêtre d'images-clés (Bundle Adjustment) :

$$\arg \min_{P_i, X_j} \left(\sum_i \sum_j (x_{ij} - \pi_i(X_j))^2 \right)$$

- ✓ Suivi image à image par KLT
- ✓ Recherche multi-vues des points occultés par KLT sur fenêtre glissante
- ✓ **UW-VO – Méthode d'odométrie visuelle monoculaire basée image-clés avec Bundle Adjustment fenêtré**

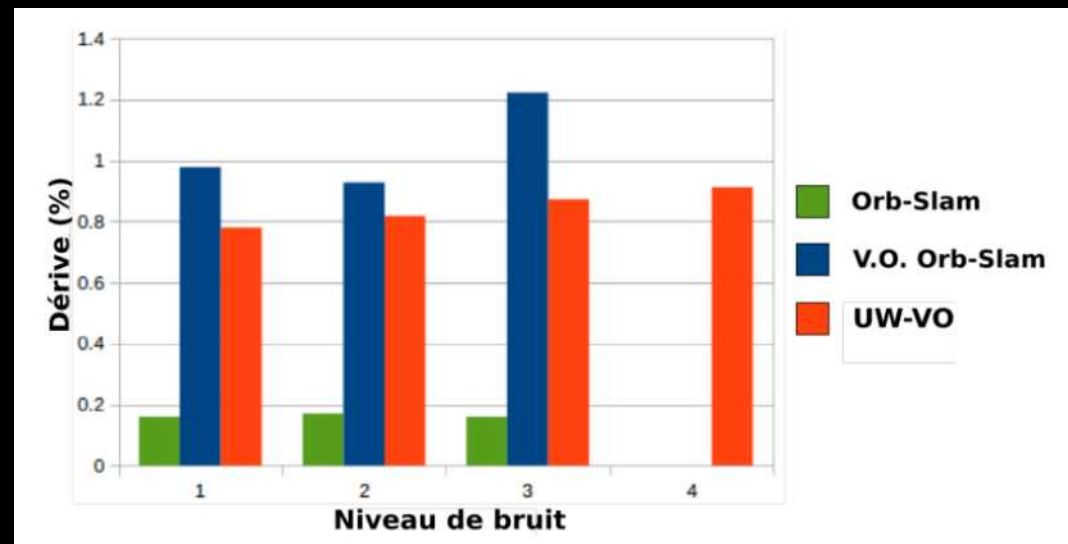
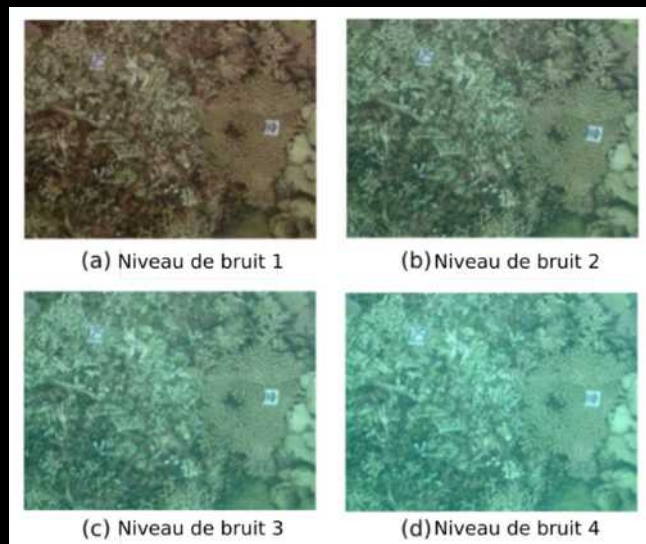
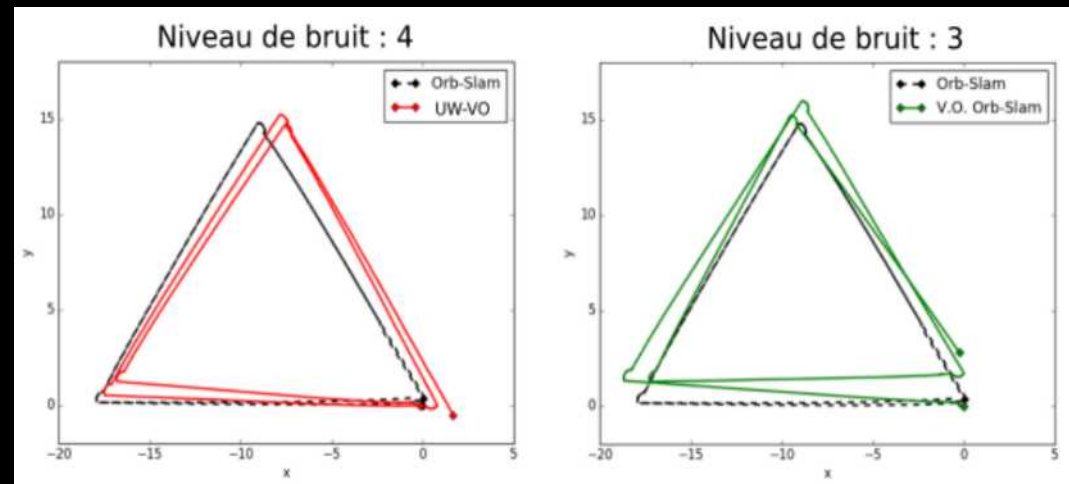
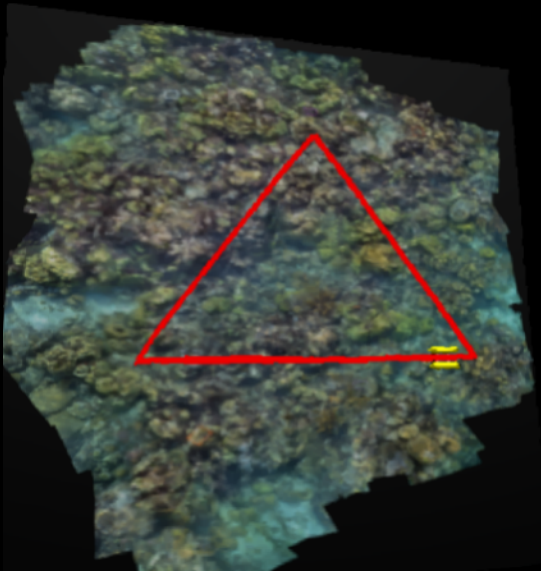
Implémentation

- Algorithme développé en C++
- Bibliothèques utilisées :
 - OpenCV
 - OpenGV¹
 - g2o²
- Threads :
 - Thread tracking + estimation de pose
 - Thread d'optimisation
- Temps de calcul : ~25ms / image (i5 @2.20GHz – 8Gb RAM)

1: L. Kneip, D. Scaramuzza, and R. Siegwart. A novel parametrization of the perspective-three-point problem for a direct computation of absolute camera position and orientation. In CVPR, 2011.

2: Kümmerle, R., Grisetti, G., Strasdat, H., Konolige, K., Burgard, W., 2011. g2o: A general framework for graph optimization. In: IEEE International Conference on Robotics and Automation (ICRA).

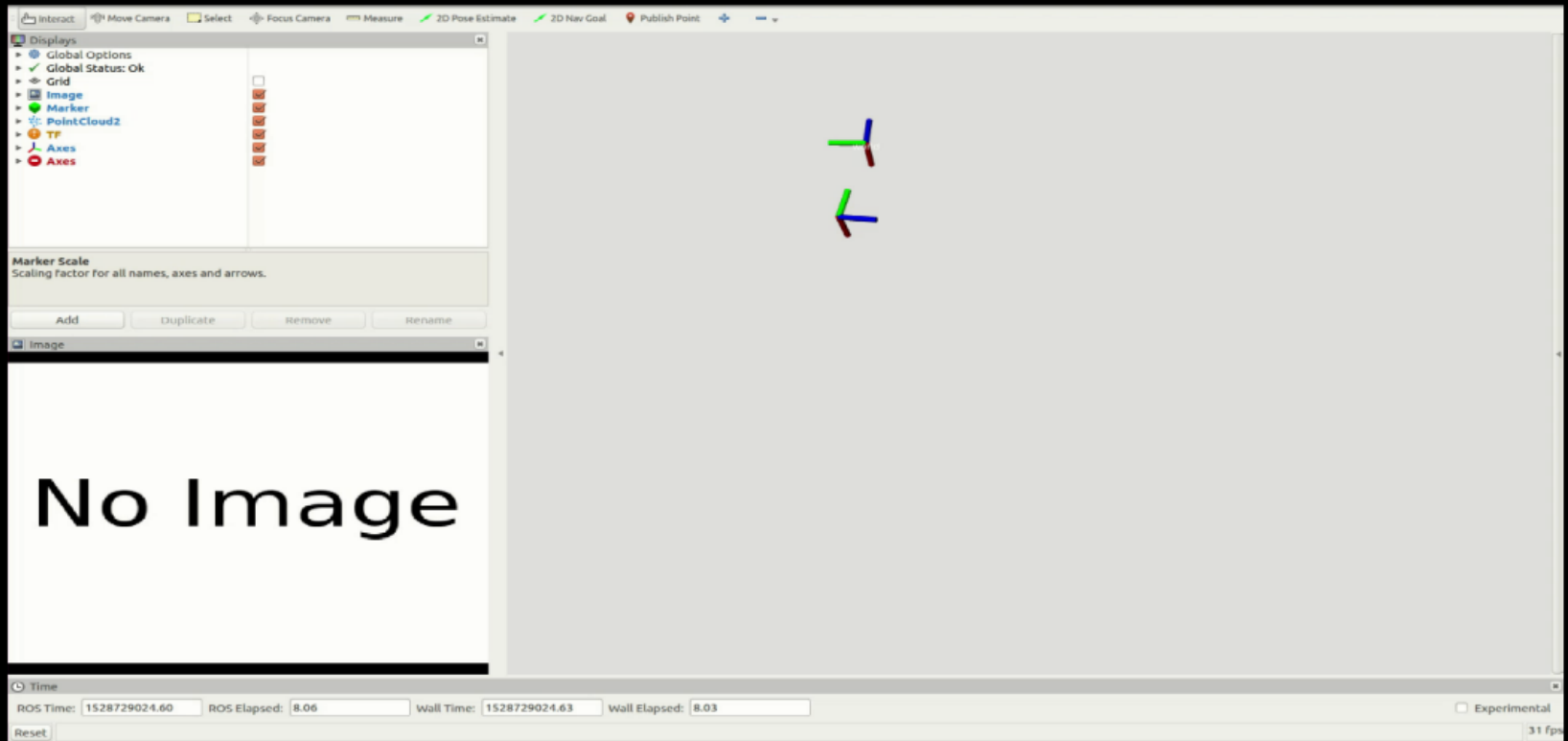
Résultats sur données simulées¹



Résultats sur données réelles

Corse 2016, Epave antique, 500 mètres de profondeur

UW-VO

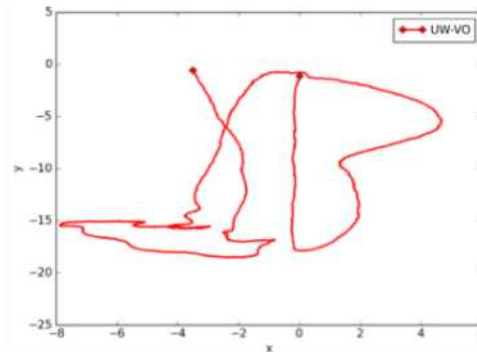


Résultats sur données réelles

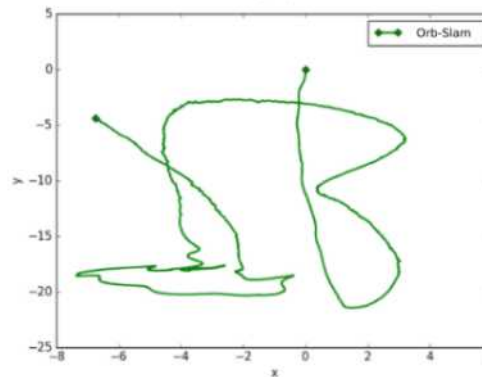
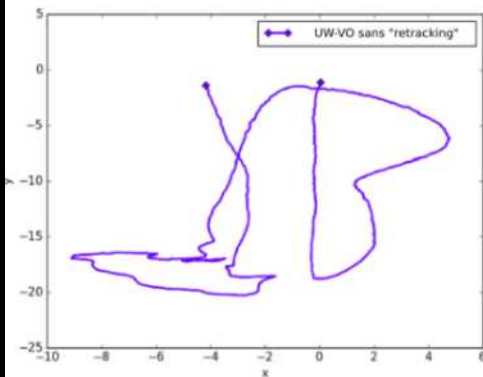
Corse 2016, Epave antique, 500 mètres de profondeur



(a)



(b)

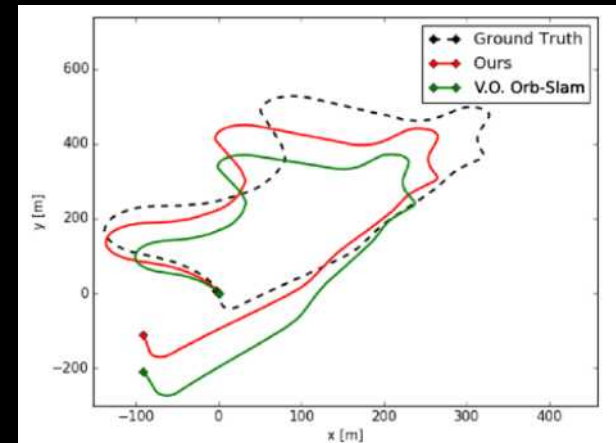


Dérive (%)		
UW-VO	UW-VO sans retracking	ORB-SLAM
1.428	1.734	3.312

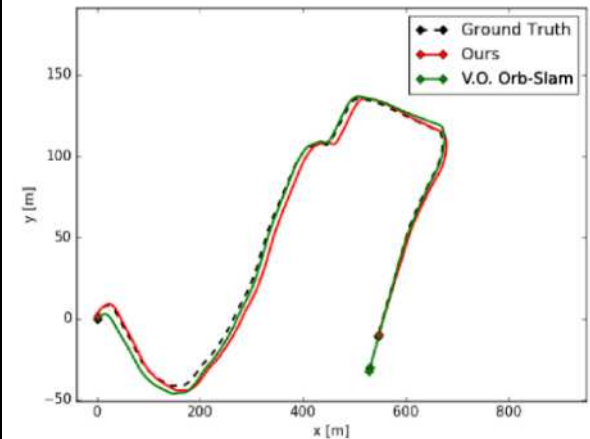
Résultats sur données réelles

KITTI¹, Séquences en milieux urbains

Séq. - Length	RMSE (m)		
	ORB-SLAM	V.O. ORB-SLAM	UW-VO
00 - 3724 m	6.68	104.89	130.4
02 - 5067 m	21.75	44.71	185.49
03 - 561 m	1.59	1.78	3.45
04 - 394 m	1.79	0.63	1.204
05 - 2206 m	8.23	44.69	65.49
06 - 1233 m	14.68	52.75	52.54
07 - 695 m	3.36	17.98	18.75
08 - 3223 m	46.58	56.45	61.02
09 - 1705 m	7.62	60.71	53.33
10 - 920 m	8.68	8.25	6.50



(a) Kitti 09



(b) Kitti 10

Conclusion

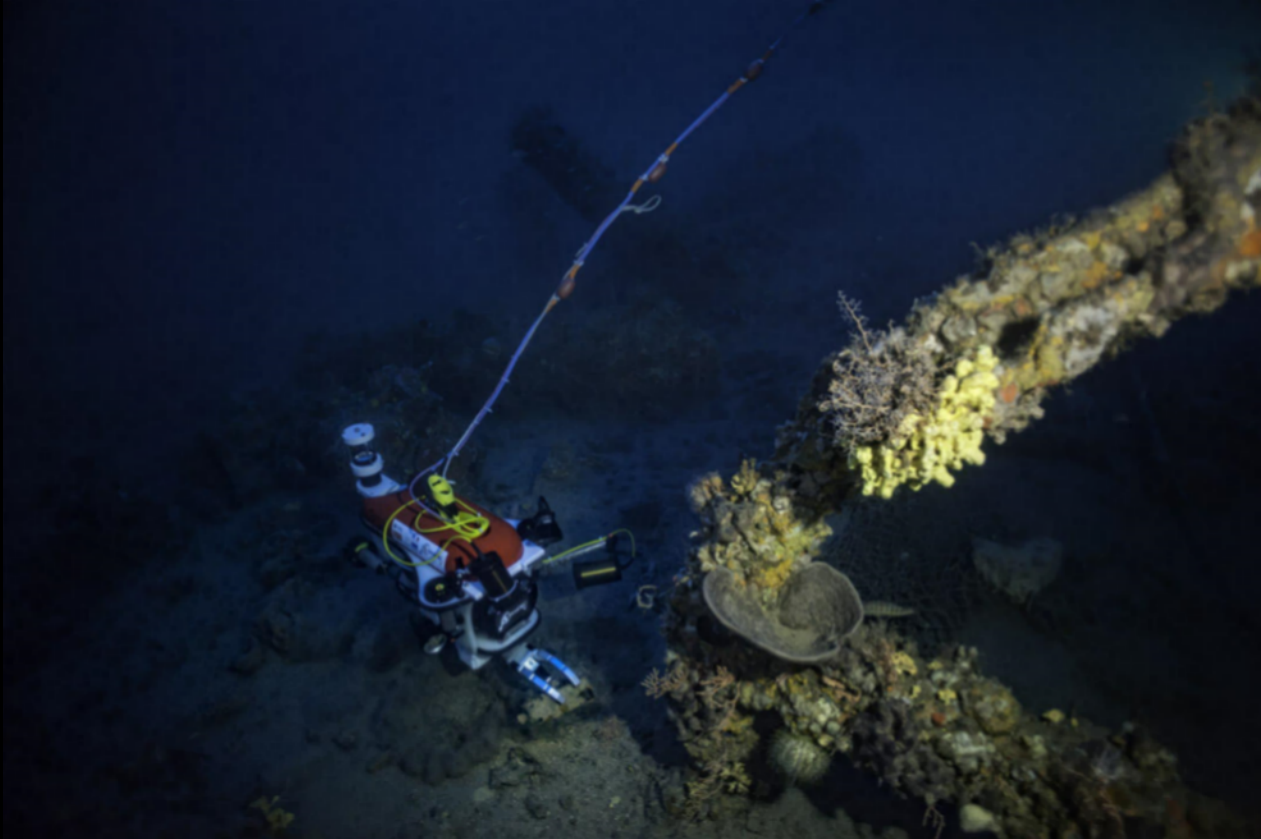
Contributions :

- Evaluation des performances des méthodes de suivi de points sur images sous-marines
- Développement de UW-VO :
 - algorithme d'odométrie visuelle robuste aux dégradations visuelles en environnement sous-marin

Perspectives :

- Fusion des mesures inertielles et de profondeur
- Détection des fermetures de boucle
- Densification des reconstructions 3D

Questions ?



Contact

Maxime Ferrera - *ONERA / LIRMM*

maxime.ferrera@onera.fr