

# LOCALISATION VISUELLE, INVARIANCE À LONG TERME ET MULTIMODALITÉ

## SÉMINAIRE DU LABORATOIRE IBISC

Fabien Bonardi  
fabien.bonardi@univ-evry.fr

Laboratoire IBISC, équipe SIAM

Jeudi 11 avril 2019



université  
PARIS-SACLAY

## ① CONTEXTE DES RECHERCHES

Robotique mobile et navigation autonome

Systèmes de vision multimodaux et mémoire visuelle

## ② PROPOSITION D'UN DESCRIPTEUR PONCTUEL MULTIMODAL

Méthodes de description des images

Résultats expérimentaux

## ③ PRISE EN COMPTE DE LA COHÉRENCE TEMPORELLE

Filtrage bayésien discret et localisation initiale

Filtrage particulière et suivi visuel de trajectoire

## ① CONTEXTE DES RECHERCHES

Robotique mobile et navigation autonome

Systèmes de vision multimodaux et mémoire visuelle

## ② PROPOSITION D'UN DESCRIPTEUR PONCTUEL MULTIMODAL

Méthodes de description des images

Résultats expérimentaux

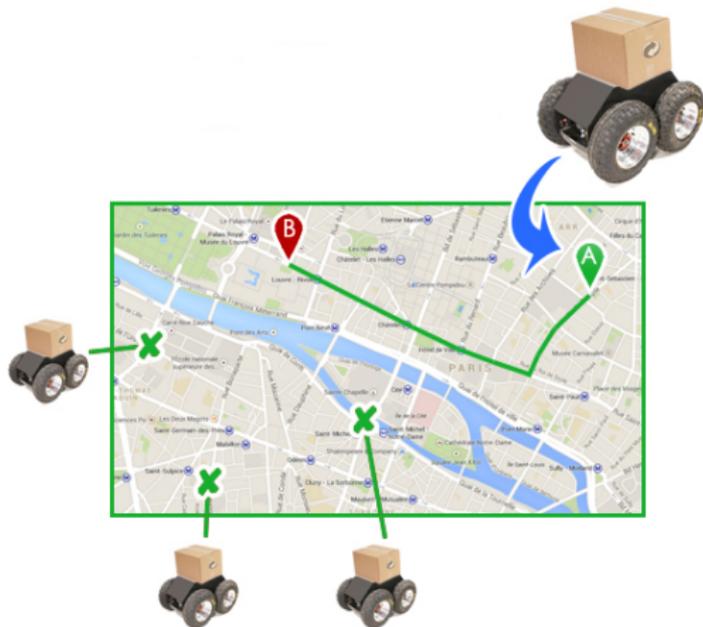
## ③ PRISE EN COMPTE DE LA COHÉRENCE TEMPORELLE

Filtrage bayésien discret et localisation initiale

Filtrage particulière et suivi visuel de trajectoire

La tâche de la navigation soulève 3 questions :

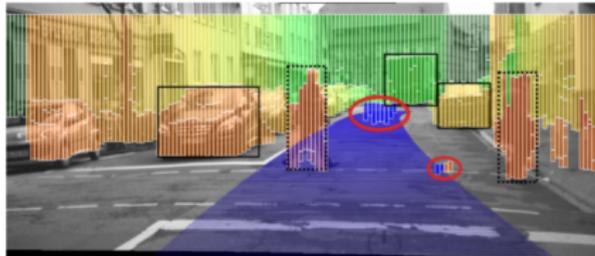
- Où suis-je ? → *localisation*
- Quelle est ma destination ? *cartographie*
- Comment y aller ? → *planification de trajectoire*



# UN SYSTÈME AU SEIN D'UN ENVIRONNEMENT DYNAMIQUE

## ÉLÉMENTS DYNAMIQUES À COURT-TERME

- Autres véhicules, piétons, cyclistes, *etc*



- Infrastructure routière



- Variations d'illuminations ponctuelles de l'environnement

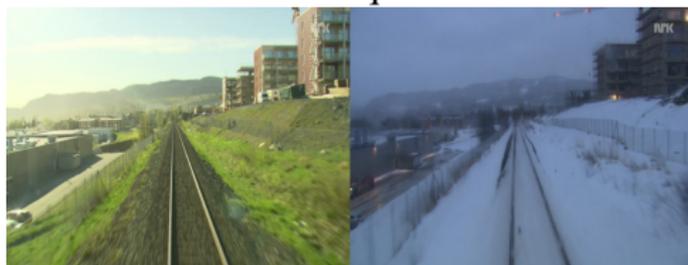
# UN SYSTÈME AU SEIN D'UN ENVIRONNEMENT DYNAMIQUE

## ÉLÉMENTS DYNAMIQUES À LONG-TERME

- Changements d'illumination dus aux heures de passage, ombres portées, navigation de nuit



- Changements de l'apparence de l'environnement dus aux variations saisonnales et climatiques



# LE VÉHICULE AU SEIN D'UN ENVIRONNEMENT DYNAMIQUE

CONSÉQUENCES : PRÉPARATIONS LONGUES ET LIÉES AUX CAPTEURS.

- Extraire de l'information de l'environnement invariante <sup>1</sup>



- Prédire les variations attendues et agir en conséquence <sup>2</sup>



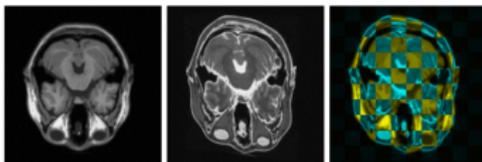
1. Will MADDERN et al. “Illumination invariant imaging : Applications in robust vision-based localisation, mapping and classification for autonomous vehicles”. *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA) Visual Place Recognition in Changing Environments Workshop*. 2014. **print**.

À l'aide de capteurs basés vision, proposer une méthode de cartographie représentant l'environnement du véhicule qui soit indépendante des caractéristiques (géométriques, spectrales, *etc*) du capteur.

# VISION MULTIMODALE

ASSOCIER DES IMAGES DE DIFFÉRENTES SENSIBILITÉS SPECTRALES

- Imagerie médicale<sup>3</sup>



- Imagerie aérienne multispectrale<sup>4</sup>



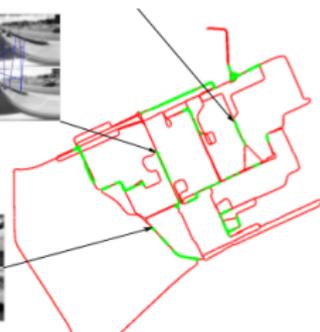
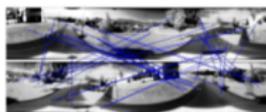
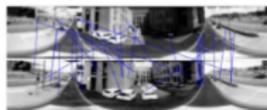
- Automatique et suivi différentiel<sup>5</sup>



3. Almar KLEIN et al. “Multimodal image registration by edge attraction and regularization using a B-spline grid”. *SPIE Medical Imaging. International Society for Optics et Photonics. 2011. 796220–796220. print.*

4. Saeed YAHYANEJAD et Bernhard RINNER. “A fast and mobile system for registration of low-altitude visual and thermal aerial images using multiple small-scale UAVs”. *ISPRS Journal of Photogrammetry and Remote Sensing* 104 (2015) : 189–202. **print.**

5. Amaury DAME et E. MARCHAND. “Une approche unifiée reposant sur l’information mutuelle pour l’asservissement visuel et le suivi différentiel”. *Proceedings of the 18e congrès francophone AEPF-AFIA Reconnaissance des Formes et Intelligence Artificielle (RFIA)*. Caen



- Fermeture de boucle<sup>6</sup>
- « Robot kidnapping »<sup>7</sup>
- Approches « Mid-Level Features »<sup>8</sup>

6. Alexandre CHAPOULIE. “Contributions aux méthodes de détection visuelle de fermeture de boucle et de segmentation topologique de l’environnement.” *Thèse de doct. Université Nice Sophia Antipolis, 2012. print.*

7. Sebastian THRUN, Wolfram BURGARD, Dieter FOX et al. *Probabilistic robotics. T. 1. MIT press Cambridge, 2005. print.*

8. Hemanth KORRAPATI. “Fermeture de boucle pour la cartographie topologique et la navigation avec des images omnidirectionnelles”. *Thèse de doct. Université Blaise Pascal-Clermont-Ferrand II, 2013. print.*

# CONTEXTE DES TRAVAUX

## LOCALISATION VISUELLE : CONCRÈTEMENT



- On possède un ensemble d'images correspondant à une expérience antérieure ;
- Lorsque l'on rencontre le même lieu, peut-on le reconnaître. . .
- . . . malgré les évolutions de l'environnement à long terme ? (illumination de la scène, saisons, *etc*)
- . . . malgré un éventuel changement de capteur ? (résolution, optique, réponse spectrale)

## ① CONTEXTE DES RECHERCHES

Robotique mobile et navigation autonome

Systèmes de vision multimodaux et mémoire visuelle

## ② PROPOSITION D'UN DESCRIPTEUR PONCTUEL MULTIMODAL

Méthodes de description des images

Résultats expérimentaux

## ③ PRISE EN COMPTE DE LA COHÉRENCE TEMPORELLE

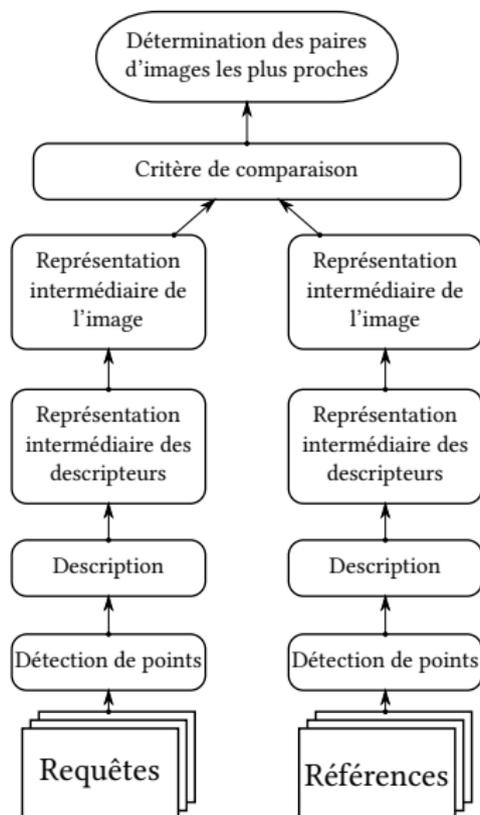
Filtrage bayésien discret et localisation initiale

Filtrage particulière et suivi visuel de trajectoire

- Description globale de l'image, de taille et caractéristiques fixes
- Descripteurs ponctuels

# MÉTHODES DE DESCRIPTION DES IMAGES

## FLUX DE TRAITEMENT PAR DESCRIPTEURS PONCTUELS



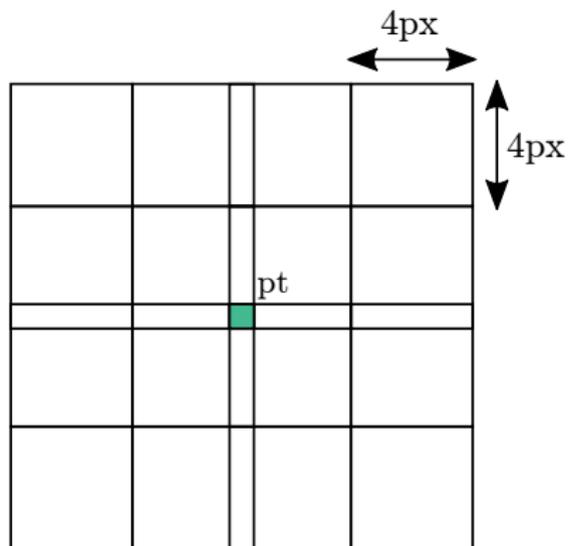


FIGURE – Motif de description du voisinage des points détectés

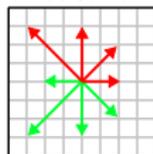
# MÉTHODES DE DESCRIPTION DES IMAGES

## PHROG : UNE NOUVELLE CARACTÉRISTIQUE

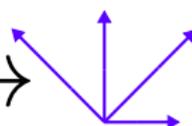


FIGURE – Inversion possible des sens des gradients d'une modalité à l'autre

Caractéristique HOG  
calculée sur un patch



HOG modifié



Pour chaque point, nous avons donc<sup>9</sup> :

- un vecteur de  $4 \times 4$  zones  $\times 4 = 64$  valeurs (128 pour SIFT),
- que nous allons extraire sur plusieurs octaves (5 maximum),
- Chaque descripteur est modifié par la suite à l'aide d'un noyau de Hellinger, plus efficace pour comparer des histogrammes.

---

9. Fabien BONARDI et al. "PHROG : A Multimodal Feature for Place Recognition". *Sensors* 17.5 (2017) : 1167. **print.**

# MÉTHODES DE DESCRIPTION DES IMAGES

DES POINTS D'INTÉRÊT VERS UNE REPRÉSENTATION COMPACTE DE L'IMAGE

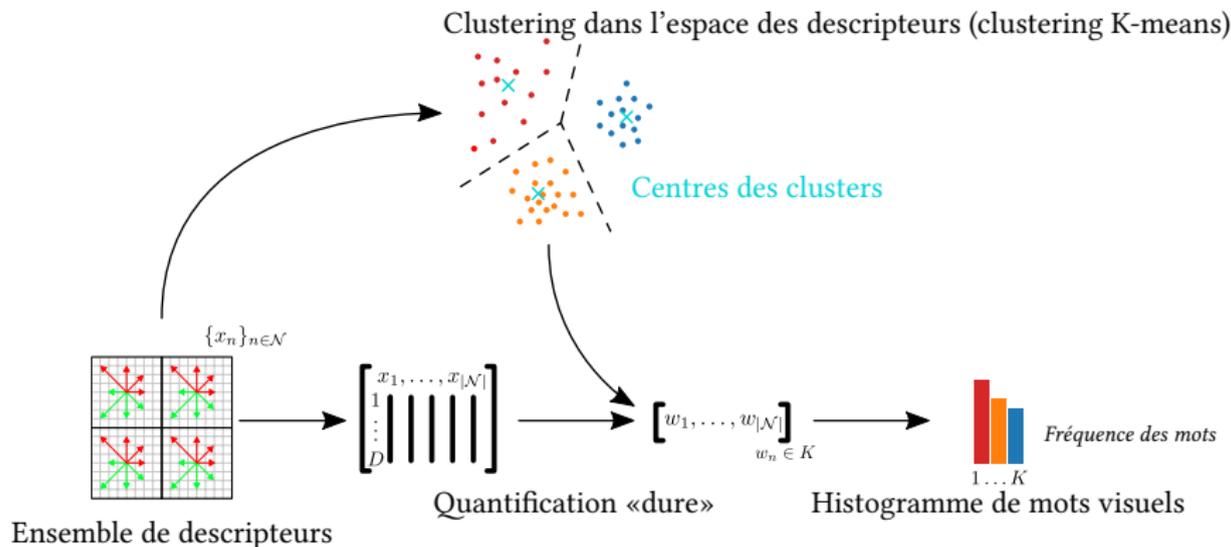
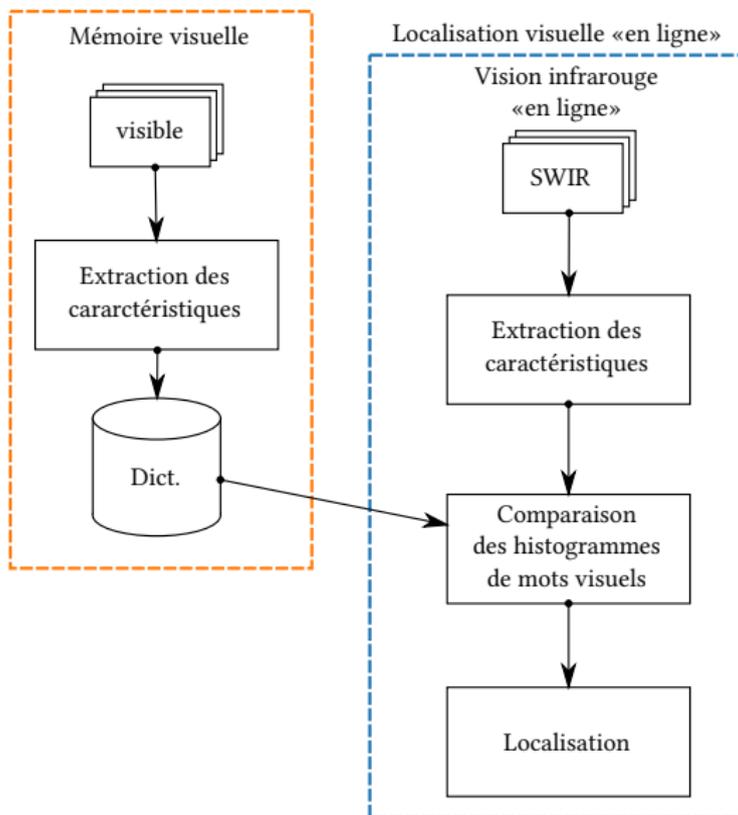


FIGURE – Principe d'une représentation des images par sacs de mots visuels

# MÉTHODES DE DESCRIPTION DES IMAGES

## VUE GLOBALE DE LA MÉTHODE D'ASSOCIATION



# RÉSULTATS EXPÉRIMENTAUX

## LES BASES D'IMAGES UTILISÉES

jeu	multimodalité	images non rectifiées	sur le long terme
EPFL	+		
Barcelone	+++		
VPRiCE		+	+
LitiSWIR	++	+	+

TABLE – Comparatif des contraintes associées à chaque jeu d'images.

# RÉSULTATS EXPÉRIMENTAUX

## DÉTERMINATION DES MATRICES DE CONFUSION

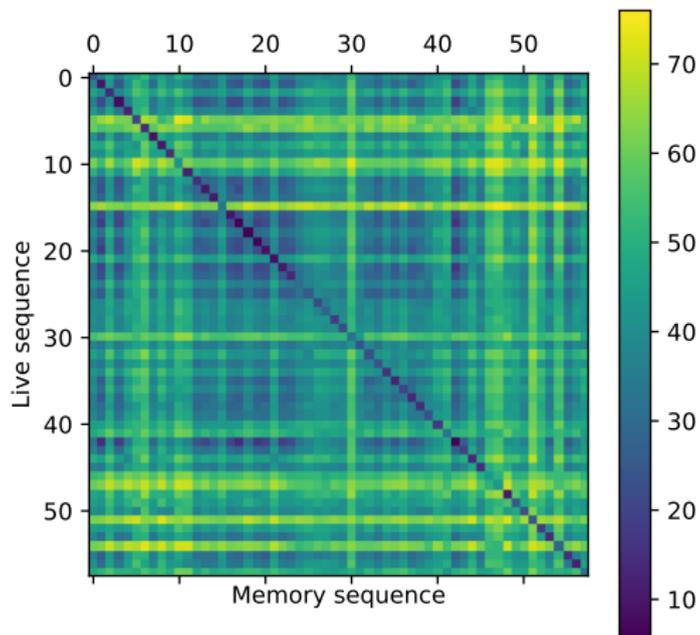


FIGURE – Matrice de confusion de deux séquences.

# RÉSULTATS EXPÉRIMENTAUX

SUR LA BASE EPFL



FIGURE – Paire d’images du jeu de l’EPFL

# RÉSULTATS EXPÉRIMENTAUX

SUR LA BASE EPFL

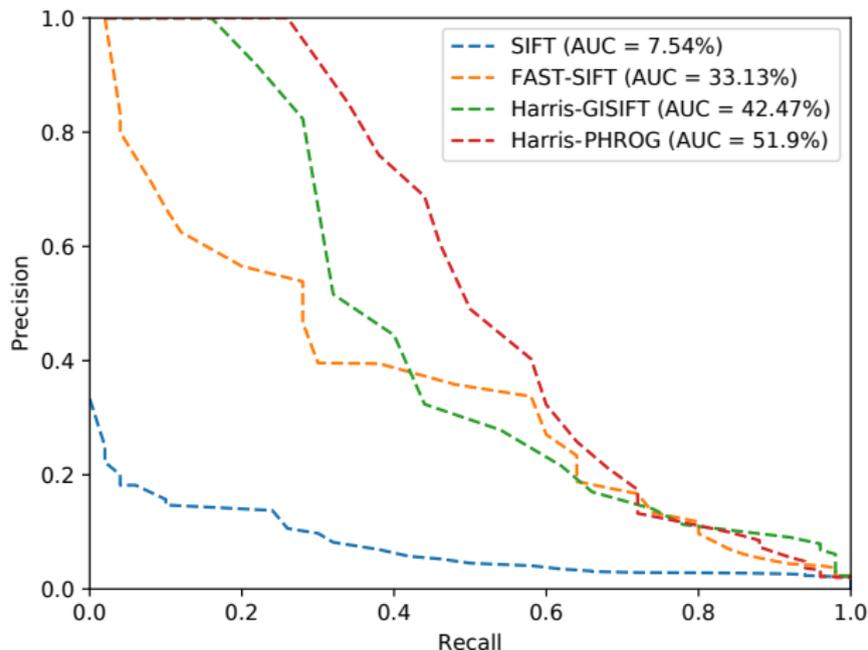


FIGURE – Courbes Précision-Rappel obtenues

# RÉSULTATS EXPÉRIMENTAUX

SUR LA BASE BARCELONE



FIGURE – Paire d'images du jeu de Barcelone

# RÉSULTATS EXPÉRIMENTAUX

SUR LA BASE BARCELONE

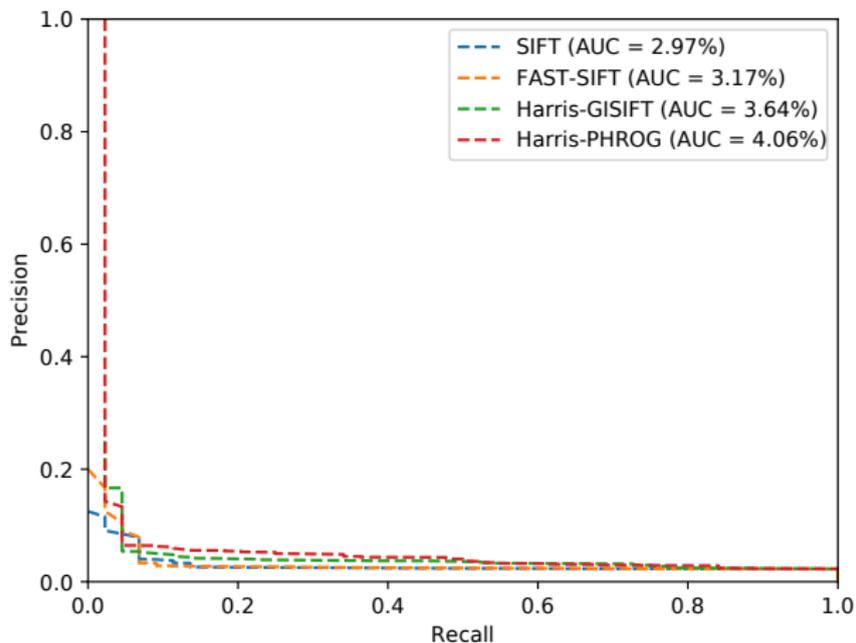


FIGURE – Courbes Précision-Rappel obtenues

# RÉSULTATS EXPÉRIMENTAUX

SUR LA BASE VPRiCE



FIGURE – Paire d’images du jeu VPRiCE

# RÉSULTATS EXPÉRIMENTAUX

SUR LA BASE VPRICE

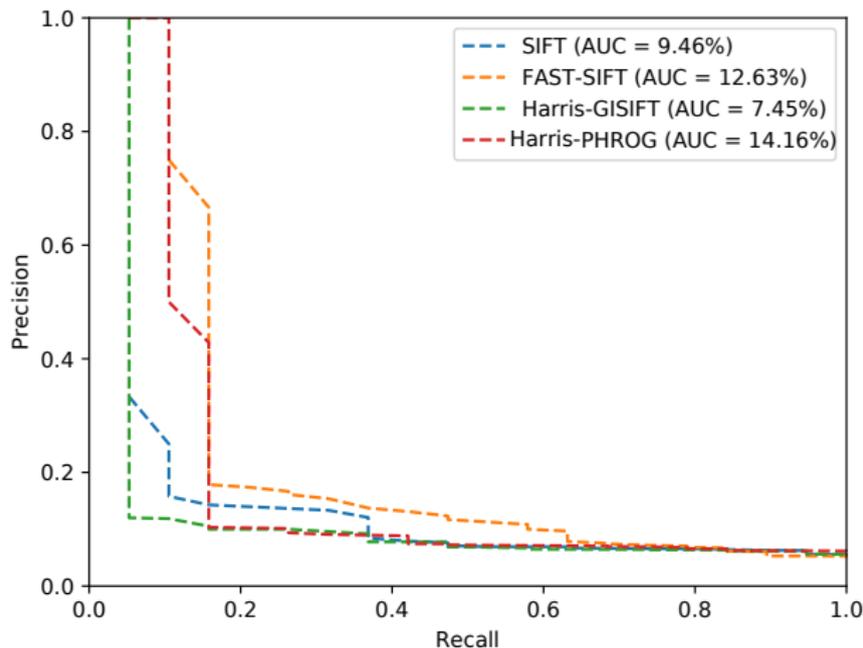


FIGURE – Courbes Précision-Rappel obtenues

# RÉSULTATS EXPÉRIMENTAUX

SUR LA BASE LITISWIR



FIGURE – Paire d’images issue de notre jeu

# RÉSULTATS EXPÉRIMENTAUX

SUR LA BASE LITISWIR

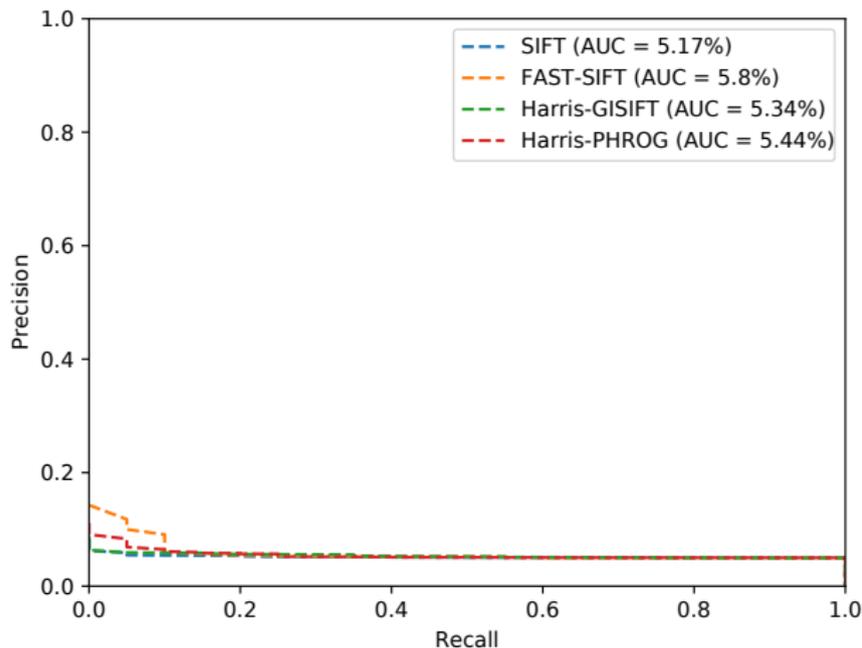


FIGURE – Courbes Précision-Rappel obtenues

## ① CONTEXTE DES RECHERCHES

Robotique mobile et navigation autonome

Systèmes de vision multimodaux et mémoire visuelle

## ② PROPOSITION D'UN DESCRIPTEUR PONCTUEL MULTIMODAL

Méthodes de description des images

Résultats expérimentaux

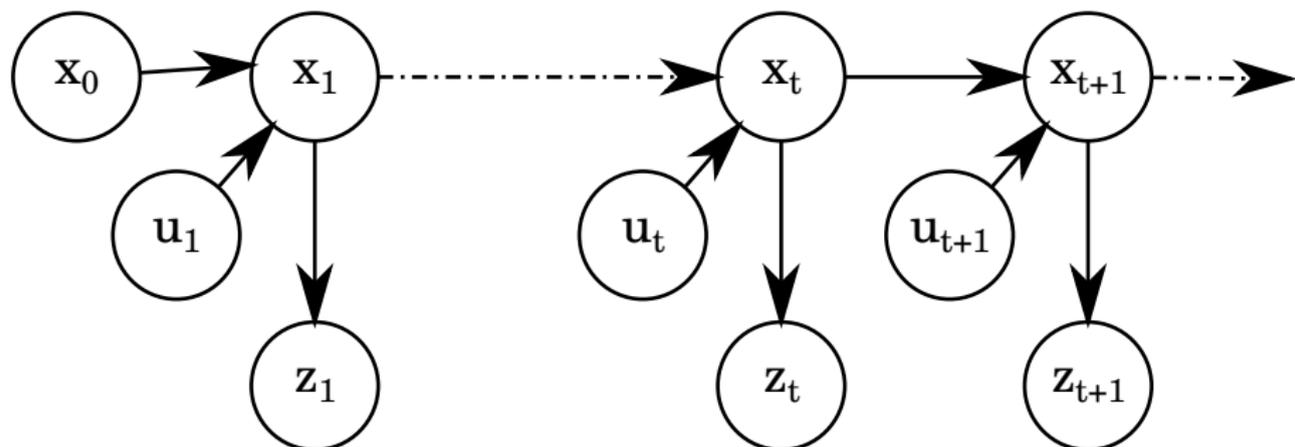
## ③ PRISE EN COMPTE DE LA COHÉRENCE TEMPORELLE

Filtrage bayésien discret et localisation initiale

Filtrage particulière et suivi visuel de trajectoire

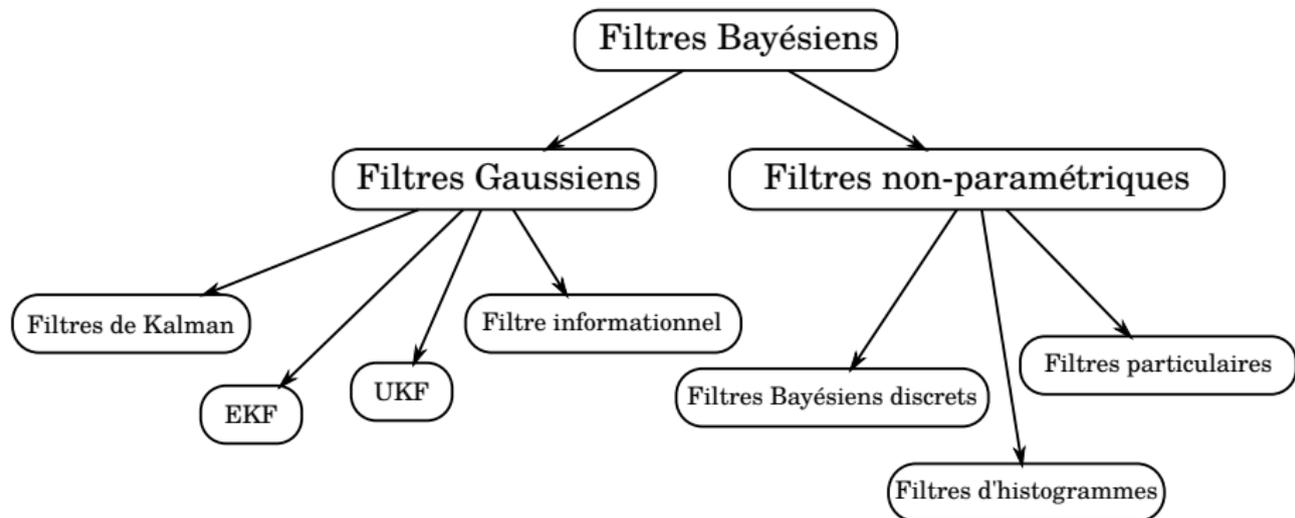
# APPROCHES PROBABILISTES

## MODÈLE DE MARKOV CACHÉ



# FILTRES BAYÉSIENS

## IMPLÉMENTATIONS DES FILTRES BAYÉSIENS



Lors de l'implémentation d'un filtre bayésien, il y a 3 hypothèses à établir :

- *probabilité de transition d'état* :  $p(x_t|x_{t-1}, u_t)$
- *probabilité de mesure* :  $p(z_t|x_t)$
- *croyance initiale* :  $bel(x_0)$  (soit  $p(x_0)$ )

Nombre fini d'états, l'algorithme du filtre bayésien discret découle directement de la définition générale :

Données en entrée :  $\{p_{k,t-1}\}$ ,  $u_t$ ,  $z_t$

À l'instant  $t$ , pour tout état  $x_t \in \Omega_X$  de probabilité  $p_{k,t}$  :

- $\bar{p}_{k,t} = \sum_i p(X_t = x_k | u_t, X_{t-1} = x_i) p_{i,t-1}$  (prédiction)
- $p_{k,t} = \eta p(z_t | X_t = x_k) \bar{p}_{k,t}$  (mise à jour par la mesure)

avec  $\eta$  tel que  $\sum_k p_{k,t} = 1$

Données en sortie :  $p_{k,t}$

On choisit comme espace d'état  $\Omega_X$  l'ensemble des index des images contenues dans la base de données préétablie. On définit alors

$bel(x_t) = p(x_t | z_{1:t}, u_{1:t}) \in \Omega_X$  la probabilité que la scène observée à l'instant  $t$  soit la même que celle de l'image d'index  $\omega$  de la base de données.

$bel(x_0) = p(x_0)$  est défini comme une densité de probabilité équitablement répartie sur l'ensemble des index de la base de données (nous n'avons pas d'a priori sur la position initiale du véhicule)

On définit les probabilités de transition d'état ainsi, pour tout  $t > 0$  et tout  $\omega \in \Omega_X$  :

$$p(X_t = \omega + 1 | X_{t-1} = \omega) = \frac{1}{3}$$

$$p(X_t = \omega - 1 | X_{t-1} = \omega) = \frac{1}{3}$$

$$p(X_t = \omega | X_{t-1} = \omega) = \frac{1}{3}$$

Autrement dit, le véhicule à l'instant  $t$  parcourt probablement la même scène qu'à l'instant  $t - 1$ , ou une scène concomitante de la base de données.

On définit les probabilités de mesure ainsi, pour tout  $t > 0$  et tout  $\omega \in \Omega_X$  :

$$p(Z_t = \omega | X_t = \omega) = 0,7$$

$$p(Z_t \neq \omega | X_t = \omega) = 0,3$$

Autrement dit, l'appariement image a 30% de chances d'être faux.

# MISE EN PRATIQUE DU FILTRE DE BAYES

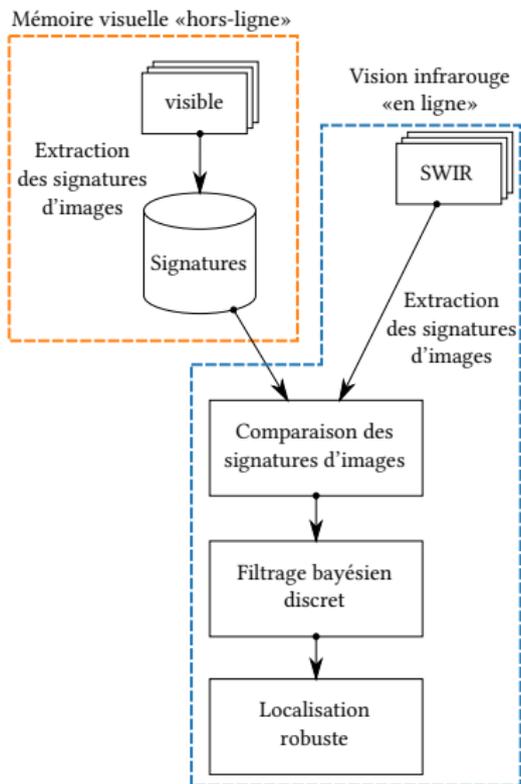


FIGURE – Schéma global de l'approche

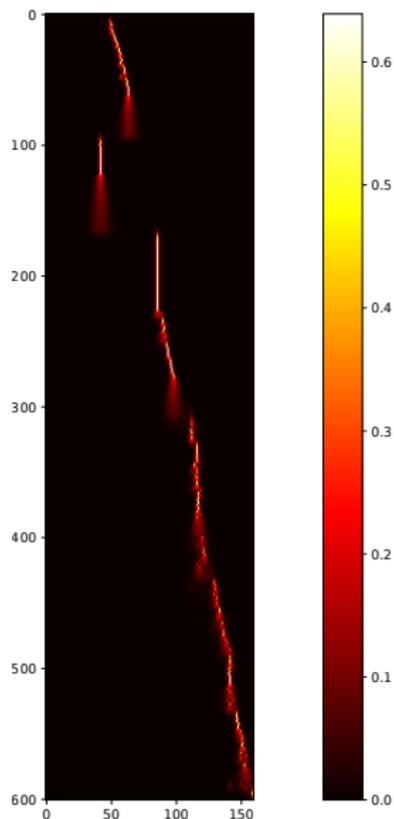
# MISE EN PRATIQUE DU FILTRE DE BAYES

SÉQUENCES UTILISÉES : IMAGES SANS FILTRE IR ET IMAGES VISIBLES



# MISE EN PRATIQUE DU FILTRE DE BAYES

RÉSULTATS : MATRICE DE CONFUSION APRÈS FILTRAGE



# MISE EN PRATIQUE DU FILTRE DE BAYES

## RÉSULTATS



**FIGURE** – Requête image noIR vers visible à long terme sans filtre (à gauche) et avec filtre bayésien (à droite)

# MISE EN PRATIQUE DU FILTRE DE BAYES

## RÉSULTATS



**FIGURE** – Requête image noIR vers visible à long terme sans filtre (à gauche) et avec filtre bayésien (à droite)

# MISE EN PRATIQUE DU FILTRE DE BAYES

## RÉSULTATS



**FIGURE** – Requête image noIR vers visible à long terme sans filtre (à gauche) et avec filtre bayésien (à droite)

# MISE EN PRATIQUE DU FILTRE DE BAYES

## RÉSULTATS



**FIGURE** – Requête image noIR vers visible à long terme sans filtre (à gauche) et avec filtre bayésien (à droite)

# MISE EN PRATIQUE DU FILTRE DE BAYES

## RÉSULTATS



**FIGURE** – Requête image noIR vers visible à long terme sans filtre (à gauche) et avec filtre bayésien (à droite)

# MISE EN PRATIQUE DU FILTRE DE BAYES

## RÉSULTATS



**FIGURE** – Requête image noIR vers visible à long terme sans filtre (à gauche) et avec filtre bayésien (à droite)

# MISE EN PRATIQUE DU FILTRE DE BAYES

## RÉSULTATS



**FIGURE** – Requête image noIR vers visible à long terme sans filtre (à gauche) et avec filtre bayésien (à droite)

# MISE EN PRATIQUE DU FILTRE DE BAYES

## RÉSULTATS



**FIGURE** – Requête image noIR vers visible à long terme sans filtre (à gauche) et avec filtre bayésien (à droite)

# MISE EN PRATIQUE DU FILTRE DE BAYES

## RÉSULTATS



**FIGURE** – Requête image noIR vers visible à long terme sans filtre (à gauche) et avec filtre bayésien (à droite)

# MISE EN PRATIQUE DU FILTRE DE BAYES

## RÉSULTATS



**FIGURE** – Requête image noIR vers visible à long terme sans filtre (à gauche) et avec filtre bayésien (à droite)

# FILTRAGE PARTICULAIRE ET SUIVI VISUEL DE TRAJECTOIRE

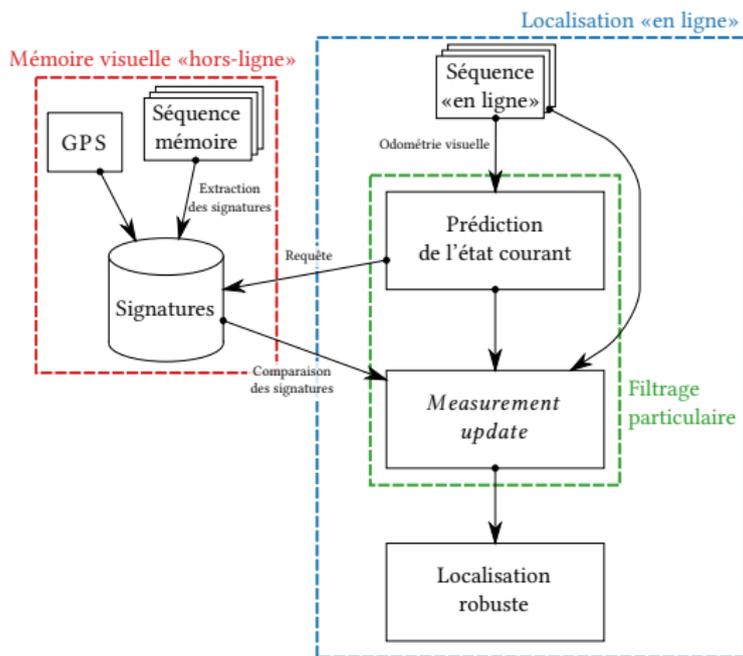


FIGURE – Schéma global de l'approche

# FILTRAGE PARTICULAIRE ET SUIVI VISUEL DE TRAJECTOIRE

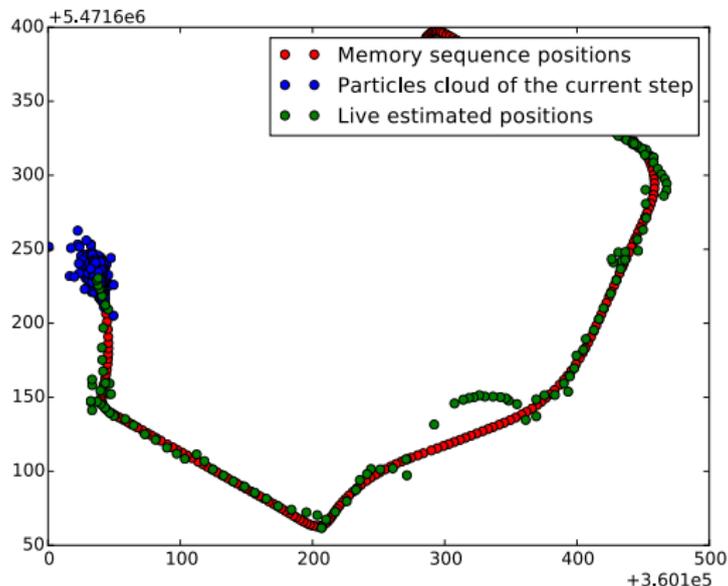
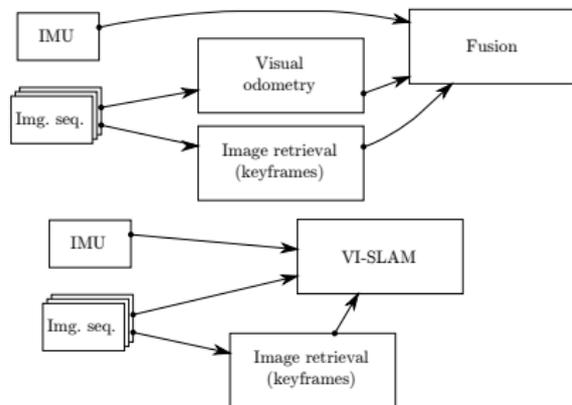
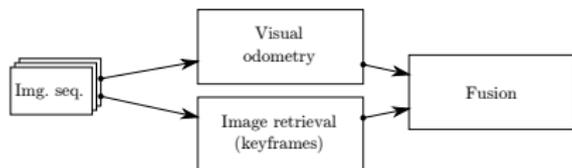


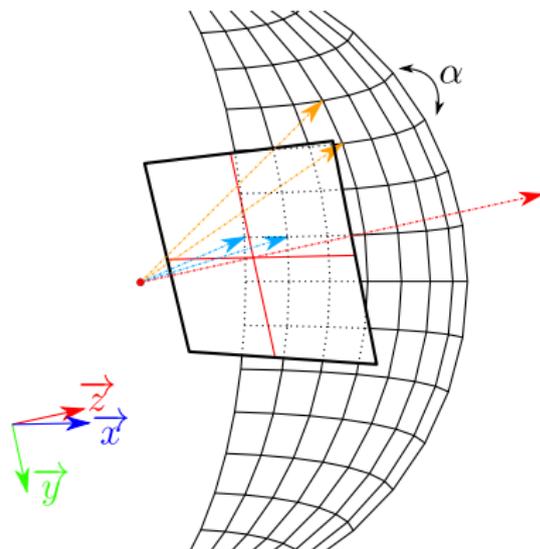
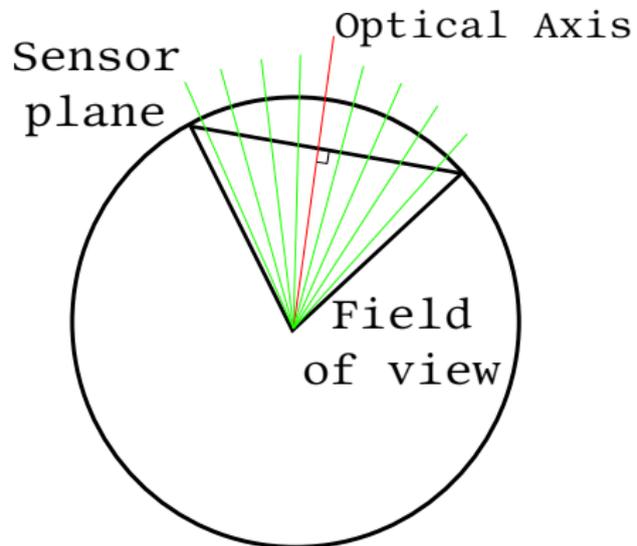
FIGURE – Résultats de l’approche avec filtre particulaire sur un trajet de courte distance

- Travaux en cours notamment au sein du projet LOCA-3D
- Fusion de données avec une centrale inertielle



# APPENDIX

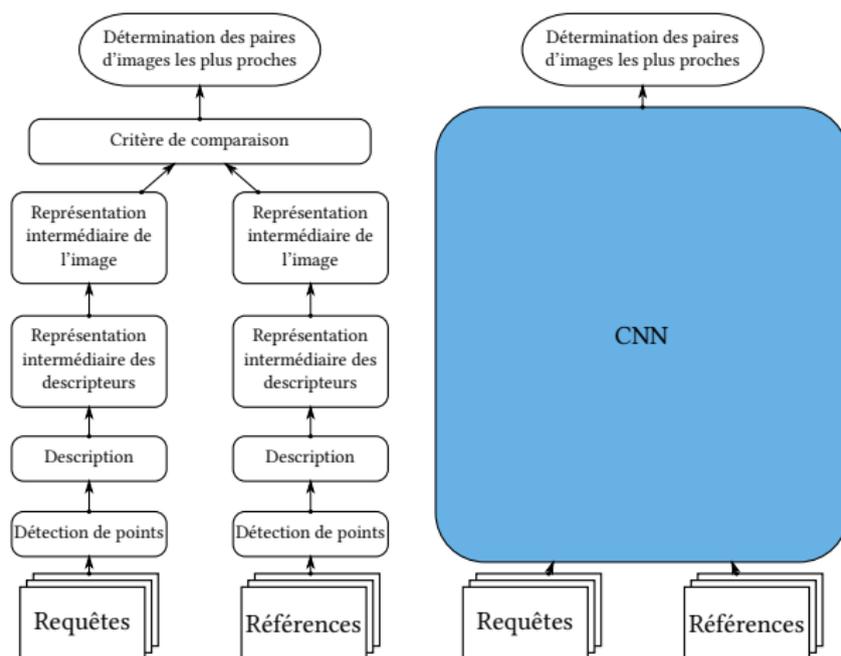
## APPROCHE PAR DESCRIPTIONS GLOBALES



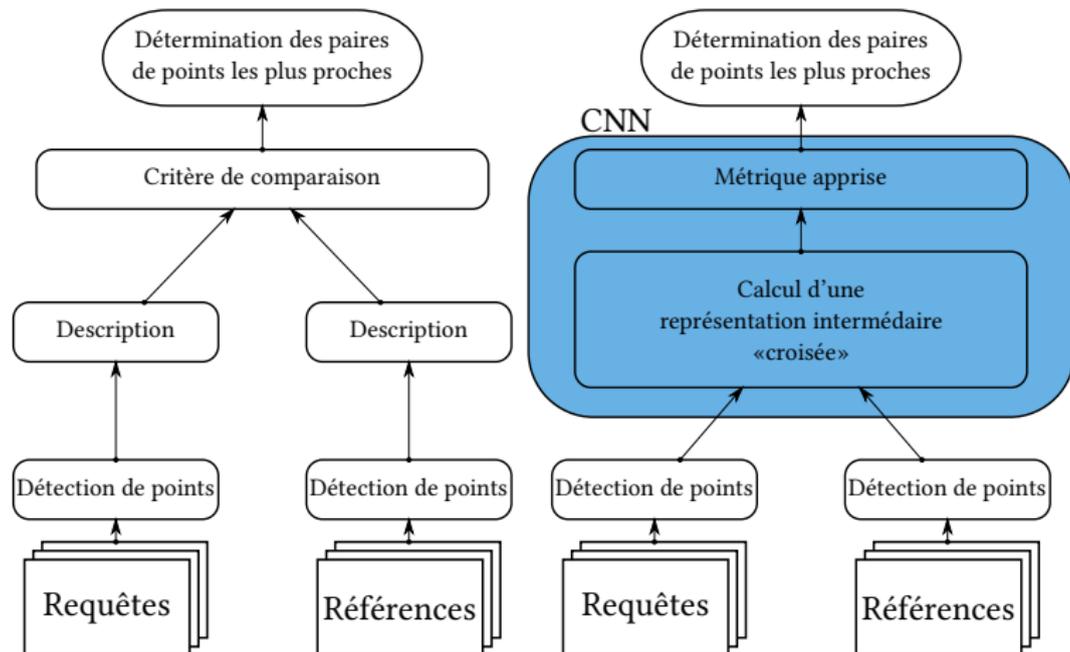
# APPENDIX

## APPROCHE PAR DESCRIPTIONS GLOBALES





10. Tobias WEYAND, Ilya KOSTRIKOV et James PHILBIN. “Planet-photo geolocation with convolutional neural networks”. *arXiv preprint :1602.05314 (2016)*. **print**.



11. Cristhian A AGUILERA et al. “Learning Cross-Spectral Similarity Measures With Deep Convolutional Neural Networks”. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*. 2016. 1–9. **print**.