

Localisation et estimation de trajectoire basées vision

Aurélien Plyer

Julien Moras, Martial Sanfourche, Guy Le Besnerais



Cours IMD - IOGS 3A-T1

Moyens

Exploiter une infrastructure de localisation externe

À base d'émission Radiofréquence



À base de vision (réseau de caméras)



Avantages

- Solutions commerciales
- Précision (Vision, GPS RTK)

Inconvénients

- Coût
- Domaine d'emploi réduit
- Interaction avec l'environnement ?

ONERA
THE FRENCH AEROSPACE LAB

Cours IMD - IOGS 3A-T2

localiser dans un environnement...

...qu'on explore

ONERA
THE FRENCH AEROSPACE LAB

localiser dans un environnement...

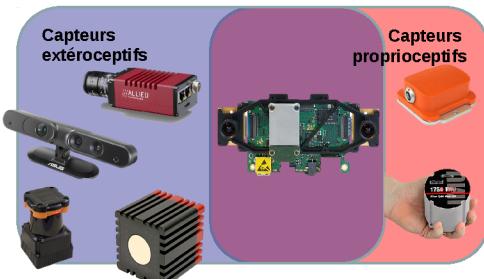
...qu'on explore

Chicken or Egg?



Moyens

Exploiter des capteurs embarqués



Avantages

- Pas d'installations coûteuses
- Versatilité
- Bas coût

Inconvénients

- Nécessite calculateurs embarqués
- Solutions pas encore industrialisées

ONERA
THE FRENCH AEROSPACE LAB

Cours IMD - IOGS 3A-T3

Localisation basée vision

Comment estimer la trajectoire d'un capteur de vision à partir des données qu'il enregistre ?

Intégration du mouvement relatif entre deux acquisitions capteur
=
Odométrie visuelle

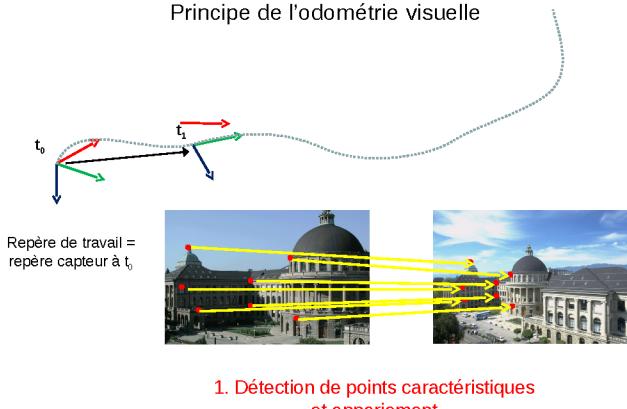
Localisation par rapport à une carte d'amers construite en ligne
=
SLAM

Localisation à l'aide d'une base de données (contenant des images, des descripteurs images, ...)
=
Relocalisation visuelle qualitative

ONERA
THE FRENCH AEROSPACE LAB

Cours IMD - IOGS 3A-T6

Principe de l'odométrie visuelle

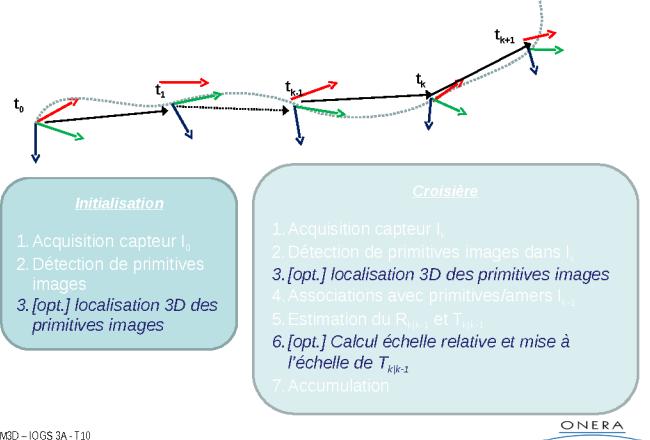


Cours IM3D – IOGS 3A - T7

Illustrations issues du tutoriel « Visual Odometry » de D. Scaramuzza

ONERA
THE FRENCH AEROSPACE LAB

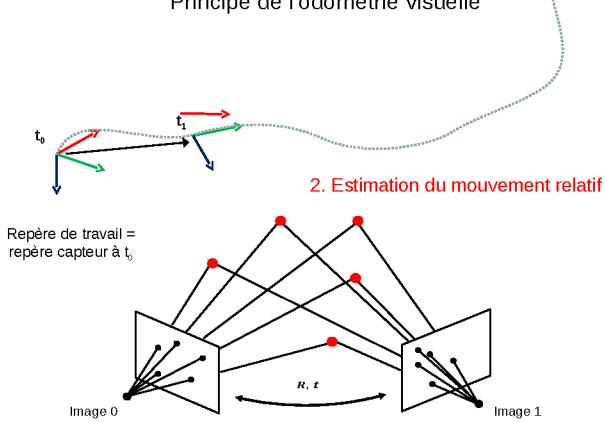
Anatomie d'un algorithme d'odométrie visuelle



Cours IM3D – IOGS 3A - T10

ONERA
THE FRENCH AEROSPACE LAB

Principe de l'odométrie visuelle

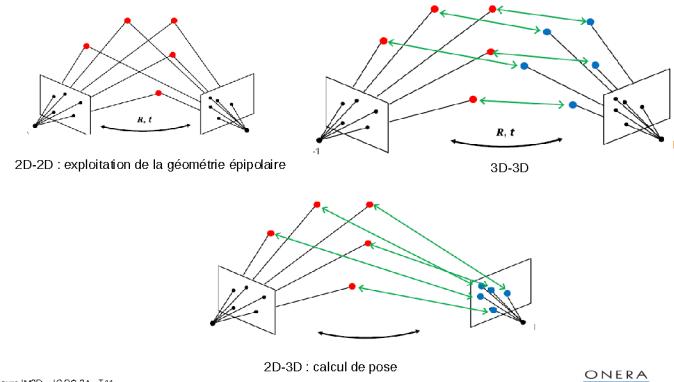


Cours IM3D – IOGS 3A - T8

Illustrations issues du tutoriel « Visual Odometry » de D. Scaramuzza

ONERA
THE FRENCH AEROSPACE LAB

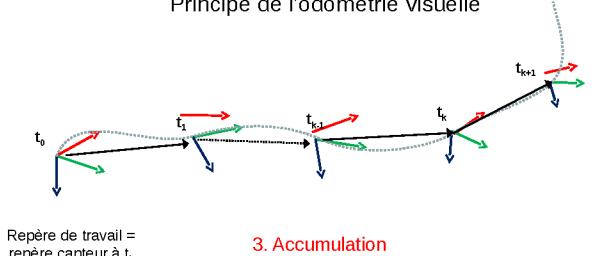
Calcul du mouvement relatif : 3 solutions



Cours IM3D – IOGS 3A - T11

ONERA
THE FRENCH AEROSPACE LAB

Principe de l'odométrie visuelle



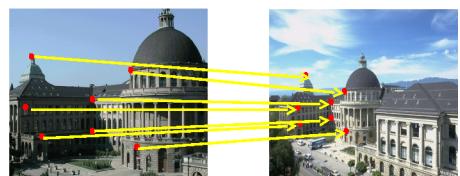
Cours IM3D – IOGS 3A - T9

ONERA
THE FRENCH AEROSPACE LAB

Estimation du mouvement 3D

• Cas 2D-2D (=pas d'information de profondeur)

- Entrées :
 - Paramètres intrinsèques de la caméra et images corrigées des distorsions
 - Points d'intérêt dans les images I_k et I_{k+1}
 - Associations entre points d'intérêt



- Pas d'estimation directe de R et T à partir de ces informations \Rightarrow estimation de la matrice essentielle et décomposition en R et T

Cours IM3D – IOGS 3A - T12

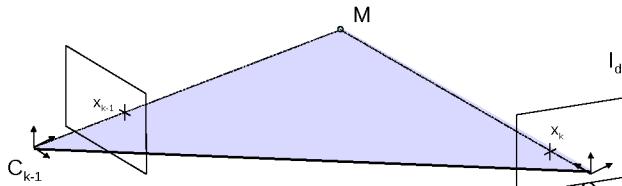
Illustrations issues du tutoriel « Visual Odometry » de D. Scaramuzza

ONERA
THE FRENCH AEROSPACE LAB

Estimation du mouvement 3D

CAS 2D-2D

- Rappel géométrie à 2



Si x_k et x_{k-1} sont des pixels homologues (= images d'un même point 3D, ici noté M) alors la relation suivante est respectée

$$\tilde{x}_k^t \mathbf{F} \tilde{x}_{k-1} = 0$$

Matrice Fondamentale

Cours IM3D - IOGS 3A - T13

ONERA
THE FRENCH AEROSPACE LAB

Estimation du mouvement 3D

CAS 2D-2D

- Lien entre matrice fondamentale (F) et matrice essentielle (E)

$$F = K_k^{-1} E K_{k-1}^{-1}$$

- Lien entre matrice essentielle et $R_{k|k-1}$ et $T_{k|k-1}$
(cf. T88 cours Multivue)

$$F = K_k^{-1} [T_{k-1|k}]^\wedge R_{k|k-1} K_{k-1}^{-1}$$

!! ATTENTION !!
Il s'agit de la position de la caméra à l'instant t_{k-1} dans le repère caméra à l'instant t_k .

Par identification :

$$E = [T_{k-1|k}]^\wedge R_{k|k-1}$$

- Estimation de la matrice E

- Algorithme des 8 points de Longuet-Higgins
 - Reviert à résoudre un système du type $A \cdot \text{vec}(E) = 0$ (par svd)
- Algorithme des 5 points de Nister (hors scope de ce cours)

D. Nister, « An efficient solution to the five-point relative pose problem », in IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol26, N°6, 2004

ONERA
THE FRENCH AEROSPACE LAB

Estimation du mouvement 3D

CAS 2D-2D

- Décomposition de la matrice Essentielle
 - Ambiguité d'échelle (norme de t non mesurable)
 - 4 solutions possibles

Méthode formelle
(Longuet-Higgins, 1981)

1. Calcul de t

$$tt^t = \frac{1}{2} \text{Trace}(EE^t) - EE^t$$

Ambiguité sur le signe de t => 2 solutions possibles

2. Calcul de R

$$t^t R = \text{Cofacteur}(E)^t - [t]_x E$$

avec

$$\text{Cofacteur}(E) = |e_1 \times e_2 \quad e_2 \times e_3 \quad e_1 \times e_3|$$

Ambiguité sur R par propagation de celle sur t => 2 solutions possibles

Décomposition SVD
1. Décomposition SVD de la matrice E et modification des valeurs singulières pour obtenir :

$$E = U \text{diag}(1,1,0) V^t$$

2. On pose $P_d = [I|0]$ et $W = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$
Alors on a 4 solution pour P_d :

$$P_d = UWV^t |u_3|$$

$$P_d = UWV^t |u_2|$$

$$P_d = UWV^t |u_1|$$

$$P_d = UWV^t |u_3|$$

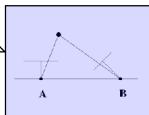
ONERA
THE FRENCH AEROSPACE LAB

Estimation du mouvement 3D

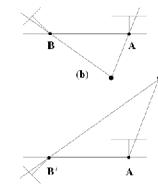
CAS 2D-2D

- Décomposition de la matrice Essentielle

Une seule solution pour laquelle le point est devant les deux caméras



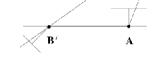
(a)



(b)



(c)



(d)

Cours IM3D - IOGS 3A - T16

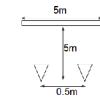
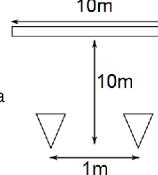
ONERA
THE FRENCH AEROSPACE LAB

Estimation du mouvement 3D

CAS 2D-2D

- Et l'échelle dans tout ça ?

Echelle absolue ne peut être calculée sans une aide extérieure



Solution : Un objet de dimension connue, une mesure de vitesse ou de déplacement par un autre système d'odométrie (à roue ou inertielle)

... mais il convient de s'assurer que l'échelle relative soit propagée correctement

Cours IM3D - IOGS 3A - T17 Grâce à des calculs de ratios de distance entre points triangulés en 3D

ONERA
THE FRENCH AEROSPACE LAB

Estimation du mouvement 3D

CAS 2D-2D

- Décomposition de la matrice Essentielle
 - Ambiguité d'échelle (norme de t non mesurable)
 - 4 solutions possibles

Méthode formelle
(Longuet-Higgins, 1981)

1. Calcul de t

$$tt^t = \frac{1}{2} \text{Trace}(EE^t) - EE^t$$

Ambiguité sur le signe de t => 2 solutions possibles

2. Calcul de R

$$t^t R = \text{Cofacteur}(E)^t - [t]_x E$$

avec

$$\text{Cofacteur}(E) = |e_1 \times e_2 \quad e_2 \times e_3 \quad e_1 \times e_3|$$

Ambiguité sur R par propagation de celle sur t => 2 solutions possibles

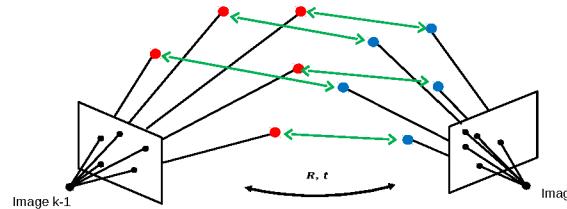
Estimation du mouvement 3D

CAS 2D-2D

- Cas 3D-3D (capteur 3D : banc stéréo, Kinect ou équivalent)

Entrées :

- Amers 3D mesurés aux instants t_{k-1} et t_k : $\{X_{k-1}(i)\}_{i \in [0, N]}$ et $\{X_k(i)\}_{i \in [0, N]}$
- Associations entre amers 3D (au nombre de $M \leq \min(N, N)$)



Estimation de R et T par minimisation aux moindres carrés du critère

Cours IM3D - IOGS 3A - T18 Illustrations inspirées du tutoriel « Visual Odometry » de D. Scaramuzza

ONERA
THE FRENCH AEROSPACE LAB

Estimation du mouvement 3D

CAS 3D-3D

- Minimisation par l'algorithme d'Umeyama (1991)

- R et T sont estimés successivement

1. Estimation de la rotation optimale $R_{k|k-1}$

- Décomposition en valeurs singulières de la matrice de covariance

$$\Sigma_{k-1,k} = \frac{1}{M} \sum_{m=1}^M (\mathbf{x}_k^m - \bar{\mathbf{x}}_k) (\mathbf{x}_{k-1}^m - \bar{\mathbf{x}}_{k-1})^T$$

$$[\mathbf{U}, \mathbf{V}, \mathbf{D}] = \text{svd}(\Sigma_{k-1,k})$$

- Appliquer la formule suivante

$$\mathbf{R}_{k-1|k} = \mathbf{U} \mathbf{S} \mathbf{V}^T$$

$$\mathbf{S} = \begin{cases} \mathbf{I}_{3 \times 3} & \text{si } \det(\Sigma_{k-1,k}) \geq 0 \\ \text{diag}(1, 1, -1) & \text{si } \det(\Sigma_{k-1,k}) < 0 \end{cases}$$

Barycentre

2. Estimation de la translation optimale $T_{k|k-1}$

$$\mathbf{T}_{k-1|k} = \bar{\mathbf{x}}_k - \mathbf{R}_{k-1|k} \cdot \bar{\mathbf{x}}_{k-1}$$

$$\mathbf{T}_{k|k-1} = -\mathbf{R}_{k-1|k}^T \cdot \mathbf{T}_{k-1|k}$$

S. Umeyama, « Least-squares estimation of transformation parameters between two point patterns », in IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 13, N°4, 1991

Cours IM3D – IOGS 3A - T19

ONERA
THE FRENCH AEROSPACE LAB

Estimation du mouvement 3D

CAS 3D-3D

A retenir

La dérive des solutions d'odométrie basée sur le recalage 3D-3D est plus importante que les autres

- En cause : incertitude de localisation des amers 3D par triangulation :

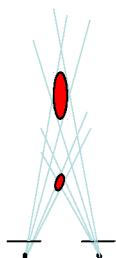
$$\sigma(z) \approx z^2 \sigma(d)$$

Exercice :

Démontrer cette propriété à partir de la relation

$$Z = \text{base} * \text{focale} / \text{disparité}$$

(cf. 178-79 du cours Multivues)



ONERA
THE FRENCH AEROSPACE LAB

Cours IM3D – IOGS 3A - T20

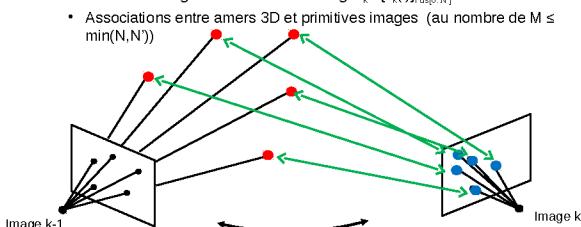
Illustration issue du tutoriel « Visual Odometry » de D. Scaramuzza

Estimation du mouvement 3D

Cas 2D-3D ou problème PnP, (Perspective from n points)

- Entrées :

- Amers 3D localisés à $\mathbf{x}_{k-1} : \{\mathbf{x}_{k-1}(i)\}_{i \in [0, N]}$
- Primitives images extraites dans l'image $\mathbf{l}_k : \{\mathbf{x}_k(i)\}_{i \in [0, N]}$
- Associations entre amers 3D et primitives images (au nombre de $M \leq \min(N, N')$)



Cours IM3D – IOGS 3A - T21

Illustrations inspirées du tutoriel « Visual Odometry » de D. Scaramuzza

ONERA
THE FRENCH AEROSPACE LAB

Estimation du mouvement 3D

CAS 2D-3D

- On recherche R et T minimisant le critère suivant :

$$\sum_{m=1}^M \|\mathbf{x}_k^m - \Pi(\mathbf{R}(\mathbf{x}_{k-1}^m - \mathbf{T}))\|^2$$



Critère non linéaire ⇒ une initialisation est nécessaire

- Concrètement :

- Estimation robuste d'une solution initiale par un algorithme de pose
 - Nombre minimal de points : 3

- Optimisation non linéaire par algorithme type gauss-Newton (Levenberg-Marquardt, Gauss-Newton)
 - Paramétrisation de la rotation par angles d'Euler ou quaternion

Cours IM3D – IOGS 3A - T22

ONERA
THE FRENCH AEROSPACE LAB

Utilisation d'une carte d'amers 3D

- Limites des approches précédentes

- La cohérence géométrique n'est assurée qu'entre deux prises de vue successives
- La localisation 3D des points est coûteuse en temps de calcul

SOLUTION :

Stocker en mémoire une carte d'amers 3D et se localiser par rapport à cette carte sur plusieurs images successives par suivi de points

- Caractéristiques de cette carte

- Un amer = un point 3D + un descripteur image + position dans les images successives
- Une mise à jour de la carte consiste à :
 - (1) supprimer les amers perdu de vue
 - (2) ajouter de nouveaux amers
 - (3) affiner la position des amers déjà observés dans un processus multi-vue

- Quand faire cette mise à jour ?

Cours IM3D – IOGS 3A - T23

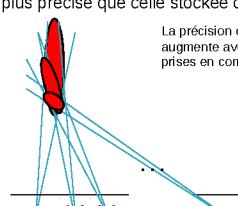
ONERA
THE FRENCH AEROSPACE LAB

Mécanisme des images-clés

- Quand mettre à jour la carte d'amers ?

- ... lorsque la précision de triangulation des points permet d'avoir une position 3D plus précise que celle stockée dans la carte

La précision de localisation par triangulation augmente avec la distance entre les images prises en compte (ligne de base)



- ... ou que cela devient nécessaire

- Car le nombre d'amers visibles décroît rapidement (mouvement de rotation)

Cours IM3D – IOGS 3A - T24

Illustration issue du tutoriel « Visual Odometry » de D. Scaramuzza

ONERA
THE FRENCH AEROSPACE LAB

eVO : un algorithme d'odométrie visuelle basée image-clé

- Un élément clé du succès d'un algorithme de VO
 - La détection au plus tôt des erreurs d'appariement (stéréo / temporel)



eVO en action

M : Gestion des amers 3D

$$S_1 = K_1$$

M

Image gauche

Détection de points caractéristiques

- ✓ DéTECTEUR de coins : Harris, FAST
- ✓ Tente d'imposer une répartition spatiale homogène
- ✓ 200 à 400 coins détectés
- ✓ Temps de traitement : 10 à 20 ms

Cours IM3D - IOGS 3A - T26

ONERA
THE FRENCH AEROSPACE LAB

eVO en action

M : Gestion des amers 3D

$$S_1 = K_1$$

M

Image gauche

Détection de points caractéristiques

Image droite

Association stéréo

disparités

Pose caméra courante

Cours IM3D - IOGS 3A - T28

ONERA
THE FRENCH AEROSPACE LAB

eVO en action

M : Gestion des amers 3D

$$S_1 = K_1$$

M

Image gauche

Détection de points caractéristiques

- ✓ DéTECTEUR de coins : Harris, FAST
- ✓ Tente d'imposer une répartition spatiale homogène
- ✓ 200 à 400 coins détectés
- ✓ Temps de traitement : 10 à 20 ms

Cours IM3D - IOGS 3A - T26

ONERA
THE FRENCH AEROSPACE LAB

eVO en action

M : Gestion des amers 3D

$$S_1 = K_1$$

M

Image gauche

Détection de points caractéristiques

Image droite

Association stéréo

disparités

Pose caméra courante

Cours IM3D - IOGS 3A - T29

ONERA
THE FRENCH AEROSPACE LAB

eVO en action

M : Gestion des amers 3D

$$S_1 = K_1$$

M

Image gauche

Association stéréo

features

- ✓ Recherche exhaustive le long des lignes épipolaires
- ✓ Score d'association : ZNCC
- ✓ Association coarse-to-fine à 2 niveaux
- ✓ Temps de calcul : ~20ms pour 250 points
- ✓ Suppression des appariements douteux par seuillage du score de corrélation

ONERA
THE FRENCH AEROSPACE LAB

eVO en action

T : Suivi des amers 3D

$$S_1 = K_1 \quad S_2$$

M

T

Image gauche

KLT: $S_{i-1} \rightarrow S_i$

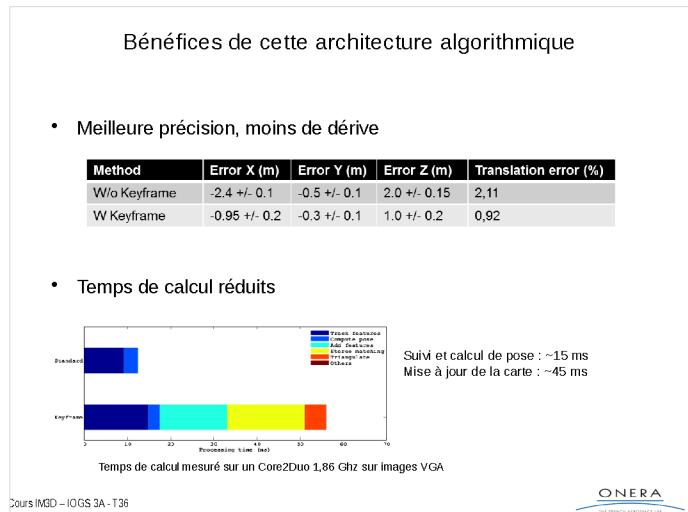
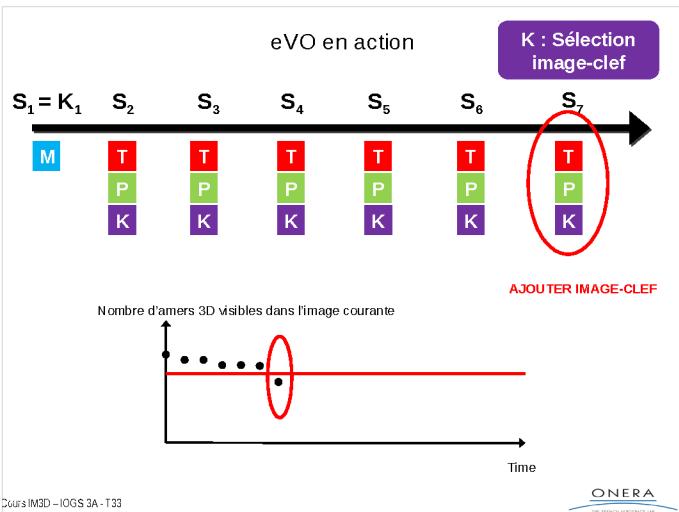
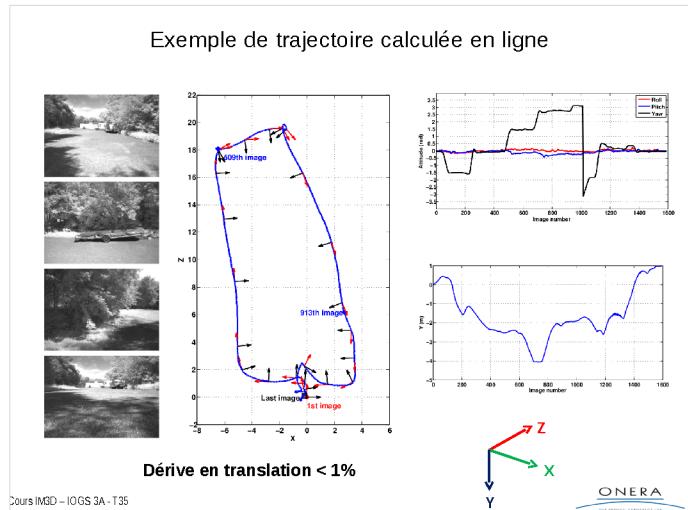
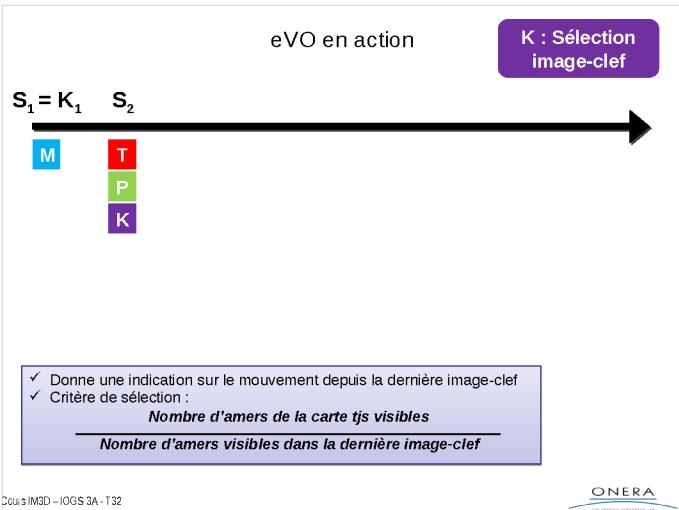
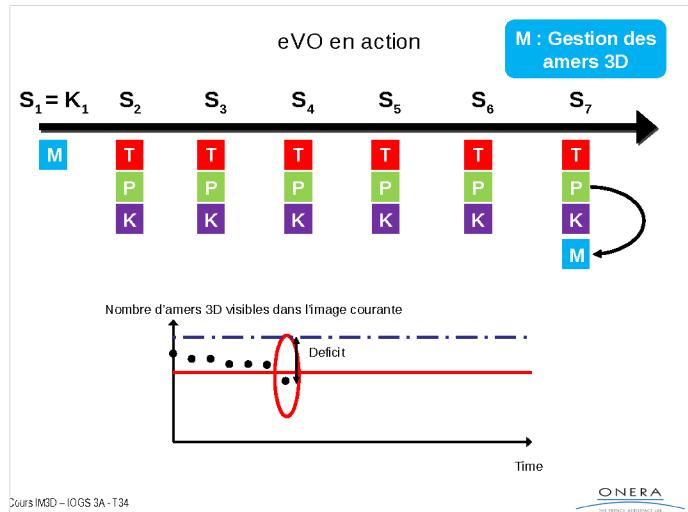
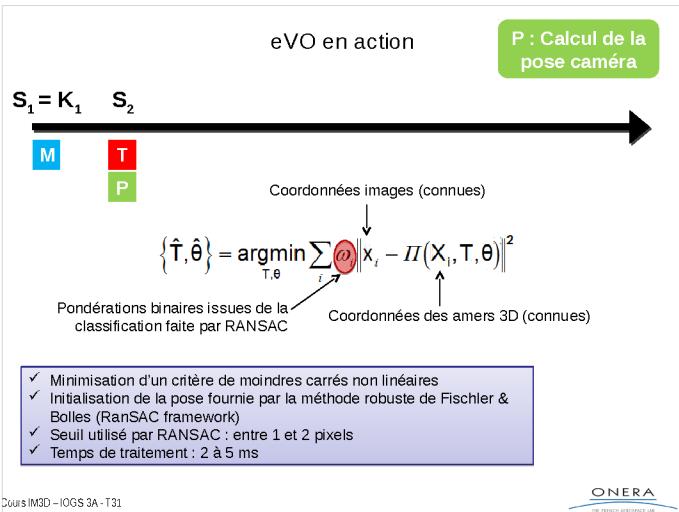
KLT: $S_i \rightarrow S_{i-1}$

Seuillage distance pixels

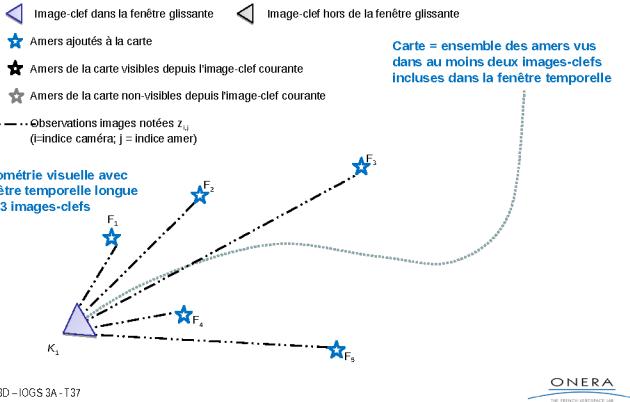
- ✓ Suivi réalisé par flot optique épart type Lucas Kanade (KLT)
- ✓ Suivi peut être initialisé par une mesure d'attitude ou un mouvement 6D
- ✓ Suppression des appariements suspects au sens d'une validation aller-retour
- ✓ Temps de traitement : ~10 ms

Cours IM3D - IOGS 3A - T27

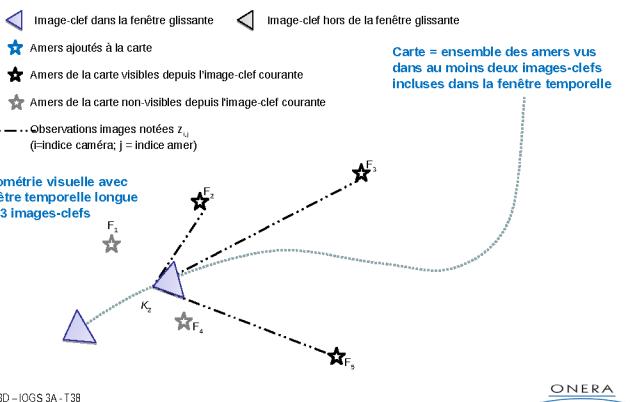
ONERA
THE FRENCH AEROSPACE LAB



Aller plus loin : Assurer une cohérence locale grâce à une fenêtre temporelle glissante

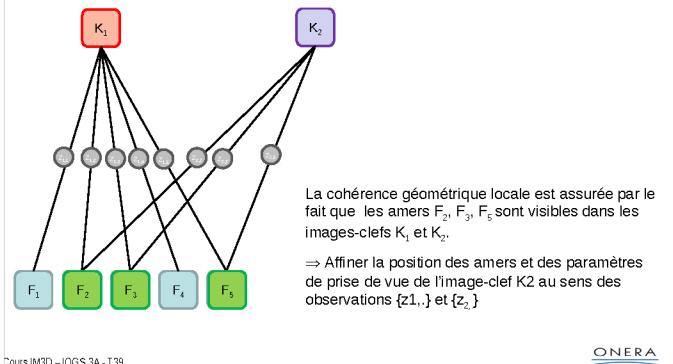


Aller plus loin : Assurer une cohérence locale grâce à une fenêtre temporelle glissante



Assurer une cohérence locale grâce à une fenêtre temporelle glissante

- Représentation sous forme d'un graphe



Ajustement de faisceaux

Formulation mathématique

- Minimisation d'un critère de reprojection des amers 3D dans les images-clés incluses dans la fenêtre temporelle

$$\varepsilon^2 = \sum_j \sum_i \omega_{i,j} \| z_{i,j} - \Pi(F_i, T_j, \theta) \|^2$$

Somme sur images-clés Somme sur amers 3D

1 si amer i visible dans image clé j
0 sinon

Procédure

- Initialisation de la pose de l'image-clé courante par calcul de pose (les autres paramètres sont déjà initialisés)
- Minimisation de l'erreur par algorithme Levenberg-Marquardt ou Région de confiance (implémentés dans la fonction lsqnonlin de Matlab)

Cours IM3D - IOGS 3A - T40

ONERA THE FRENCH AEROSPACE LAB

Ajustement de faisceaux

Problème d'optimisation de grande taille :

- 6 paramètres par image-clé
- 3 paramètres par amers

Exploiter la sparsité des données

- L'observation z_{ij} = projection de l'amer F_i sur l'image-clé K_j

$$z_{ij} = \begin{pmatrix} K_1 & K_2 & K_3 & F_1 & F_2 & F_3 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} \begin{pmatrix} \lambda_{11} & 0 & 0 & B_{11} & 0 & 0 \\ 0 & \lambda_{21} & 0 & B_{21} & 0 & 0 \\ 0 & 0 & \lambda_{31} & B_{31} & 0 & 0 \\ 0 & 0 & 0 & \lambda_{12} & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda_{22} & 0 \\ 0 & 0 & 0 & 0 & 0 & \lambda_{32} \end{pmatrix} \begin{pmatrix} U_1 & 0 & 0 & W_{11} & W_{21} & W_{31} & W_{41} \\ 0 & 0 & U_2 & 0 & W_{12} & W_{22} & W_{32} & W_{42} \\ 0 & 0 & U_3 & W_{13} & W_{23} & W_{33} & W_{43} \\ W_{11}^T & W_{21}^T & W_{31}^T & V_1 & 0 & 0 & 0 \\ W_{12}^T & W_{22}^T & W_{32}^T & 0 & V_2 & 0 & 0 \\ W_{13}^T & W_{23}^T & W_{33}^T & 0 & 0 & V_3 & 0 \\ W_{14}^T & W_{24}^T & W_{34}^T & 0 & 0 & 0 & V_4 \end{pmatrix} - \begin{pmatrix} U & W \\ W^T & V \end{pmatrix}$$

Structure de la matrice jacobienne (J) Structure de la matrice pseudo-hessienne (JJ^T)

Cours IM3D - IOGS 3A - T41

Illustrations fournies par M.I.A. Lourakis

ONERA THE FRENCH AEROSPACE LAB

Assurer une cohérence locale grâce à une fenêtre temporelle glissante

Image-clé dans la fenêtre glissante Image-clé hors de la fenêtre glissante

Amers ajoutés à la carte

Amers de la carte visibles depuis l'image-clé courante

Amers de la carte non-visibles depuis l'image-clé courante

- Observations images notées z_{ij} (i =indice caméra; j =indice amer)

Odométrie visuelle avec fenêtre temporelle longue de 3 images-clés

Cours IM3D - IOGS 3A - T42

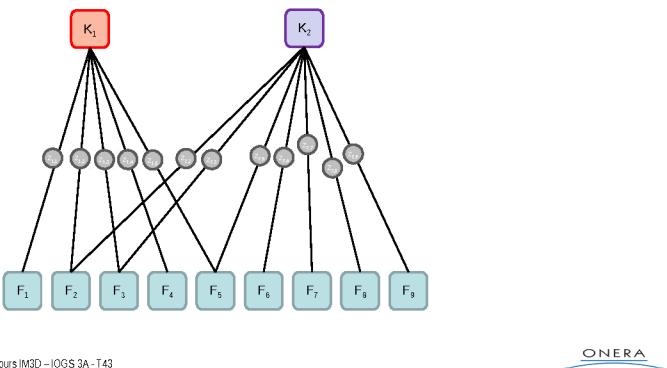
ONERA THE FRENCH AEROSPACE LAB

Cours IM3D - IOGS 3A - T42

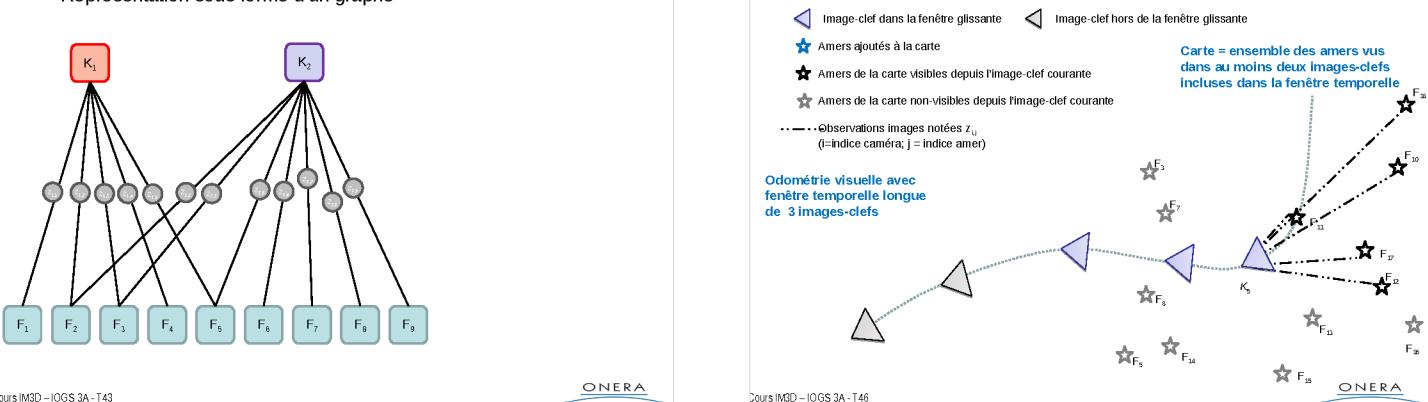
ONERA THE FRENCH AEROSPACE LAB

Assurer une cohérence locale grâce à une fenêtre temporelle glissante

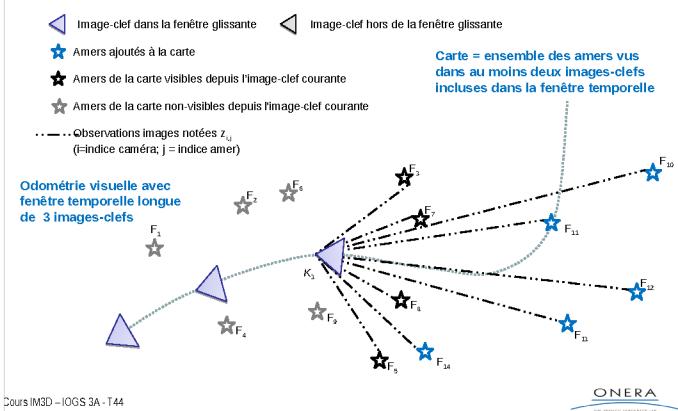
- Représentation sous forme d'un graphe



Assurer une cohérence locale grâce à une fenêtre temporelle glissante



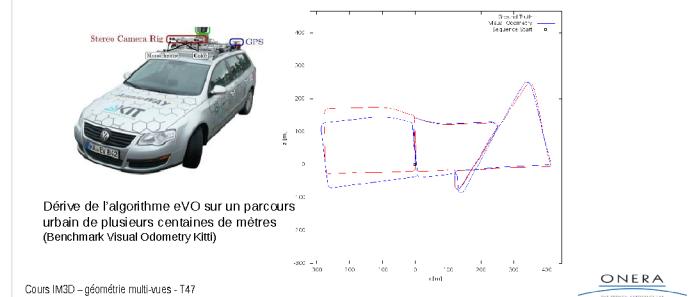
Assurer une cohérence locale grâce à une fenêtre temporelle glissante



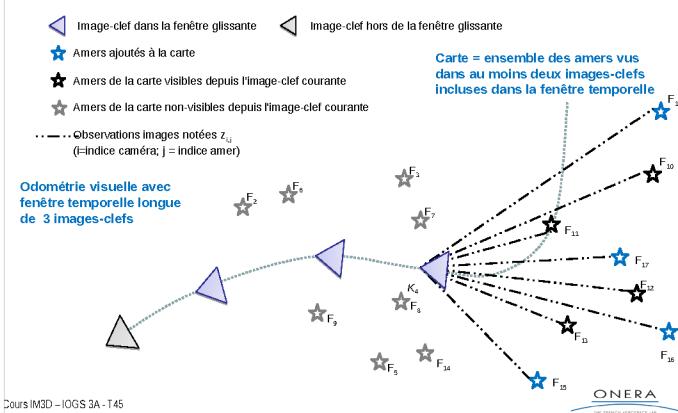
Limites de l'odométrie visuelle

L'odométrie visuelle

- Permet de reconstruire la trajectoire du capteur par intégration
- Assure une cohérence locale de la carte (de 2 à plusieurs vues)
- Ceci entraîne une dérive de la trajectoire estimée qui peut devenir très importante lors de longues opérations.



Assurer une cohérence locale grâce à une fenêtre temporelle glissante



Localisation basée vision

Comment estimer la trajectoire d'un capteur de vision à partir des données qu'il enregistre ?

Intégration du mouvement relatif entre deux acquisitions capteur
=
Odométrie visuelle

Localisation par rapport à une carte d'amers construite en ligne
=
SLAM

Localisation à l'aide d'une base de données (contenant des images, des descripteurs images, ...)
=
Relocalisation visuelle qualitative

Problème du SLAM

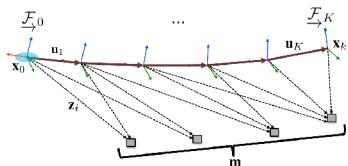
- Le problème de SLAM (Simultaneous localization and Mapping) est un problème clef dans les applications robotiques. Il s'agit d'un problème générale qui n'est pas réservé qu'aux capteurs de vision. La littérature est donc très diverse sur ce sujet.
- Il consiste à résoudre les **deux problèmes** abordés précédemment de façon **simultané**.
- Les approches utilisées pour résoudre le SLAM permettent de prendre en compte les informations provenant de **plusieurs capteurs**, souvent, de types différents. Elles réalisent alors la fusion de ces informations.
- Par exemple :
 - Capteur de vision et capteurs inertIELS pour un drone
 - Capteur lidar et capteurs de tours de roues pour un robot roulant

Cours IM3D – géométrie multi-vues - T49



Principe du SLAM

- Tous les amers sont conservés en mémoire durant toute la mission
- Leur position est remise à jour au cours du processus



- Les mesures étant bruitées, la position du robot et des amers sont connue avec des **incertitudes**. Si on ne prend pas en compte ces incertitudes, la cohérence globale de la carte ne peut être assurée.
- En pratique, on utilise souvent le cadre Bayésien pour modéliser et propager ces incertitudes.

Cours IM3D – géométrie multi-vues - T50



Localisation basée vision

Comment estimer la trajectoire d'un capteur de vision à partir des données qu'il enregistre ?

Intégration du mouvement relatif entre deux acquisitions capteur
= Odométrie visuelle

Localisation par rapport à une carte d'amers construite en ligne
= SLAM

Localisation à l'aide d'une base de données (contenant des images, des descripteurs images, ...)
= Relocalisation visuelle qualitative

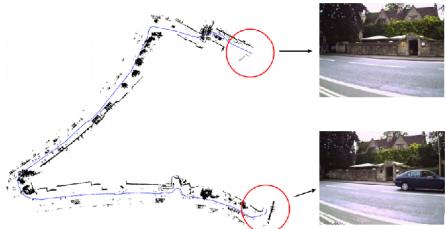
Cours IM3D – IOGS 3A - T51



Fermeture de boucle basé sur l'apparence

• Problème

- Même avec des approches de type SLAM, la trajectoire peut dériver lors de longue boucles, si on est pas capable d'associer les amers, il faudrait trouver une solution pour le recaler



• Solution

- Reconnaitre un endroit que l'on a déjà visité auparavant en se basant non plus sur la position d'amers mais sur l'apparence

Cours IM3D – géométrie multi-vues - T52



Fermeture de boucle

• Comment faire ???

- Stocker toutes les images et leur position dans une base et comparer l'image courante avec notre base d'image ???

Problème :

- Espace nécessaire
- Coût calculatoire
- Robustesse au changements
- Confondre des zone homogène qui se ressemble

- Stocker tous les points d'intérêt dans une base et comparer les points d'intérêt courant avec notre base ???

Problème :

- Les points d'intérêt seul ne sont pas discriminant (ex. les coin de fenêtres)
- La comparaison des amers va nécessiter une charge de calcule importante

Cours IM3D – géométrie multi-vues - T53



Approches Sac de mots

- Les approches « sac de mots » comme FabMap sont des approches assez performantes pour traiter ce problème.

• Phase hors ligne :

- Construction d'un **dictionnaire de « descripteurs type »**

• Phase en ligne

- Création d'une signature de l'image
 - Extraction de descripteurs sur l'image courante
 - Association avec les descripteurs du dictionnaire
 - Construction d'un vecteur qui précise quels « **descripteurs type** » sont présents dans l'image
=> **Signature de l'image**

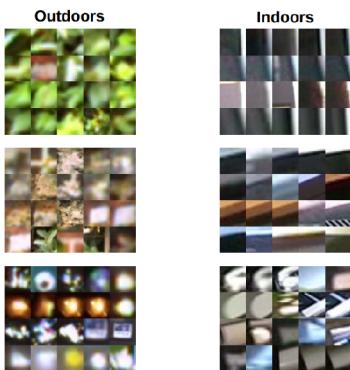
- Enregistre cette signature dans une carte en l'associant à la position que l'on a estimé par VO ou SLAM

- Comparer la signature courante aux autres signature dans la carte
 - Si on détecte une concordance, on peut recalcer la position

Cours IM3D – géométrie multi-vues - T54

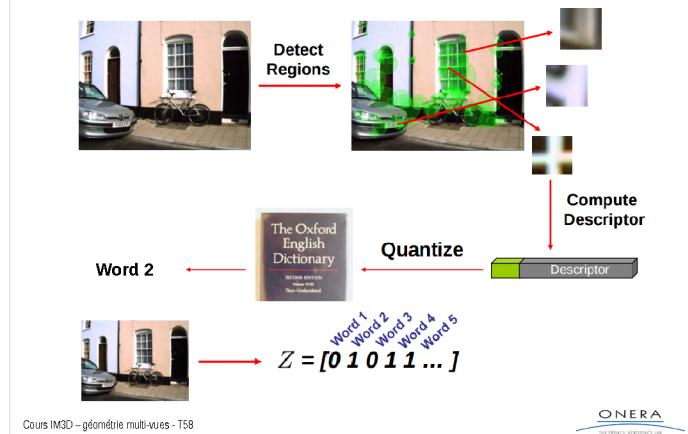


Exemple de dictionnaires



Cours IM3D – géométrie multi-vues - T55

Relocalisation



Cours IM3D – géométrie multi-vues - T58

ONERA
THE FRENCH AEROSPACE LAB

Apprentissage du dictionnaire

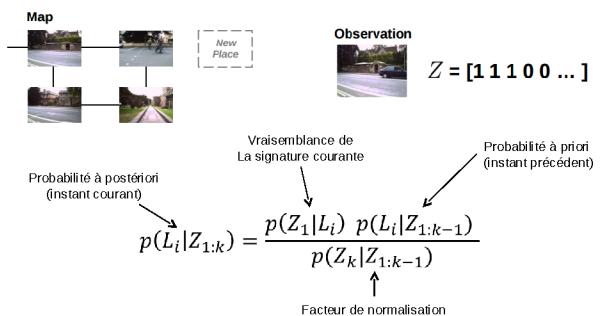
- Le dictionnaire est construit hors ligne par apprentissage sur des données collectées dans un environnement similaire à la zone d'opération.



Cours IM3D – géométrie multi-vues - T56

ONERA
THE FRENCH AEROSPACE LAB

Relocalisation

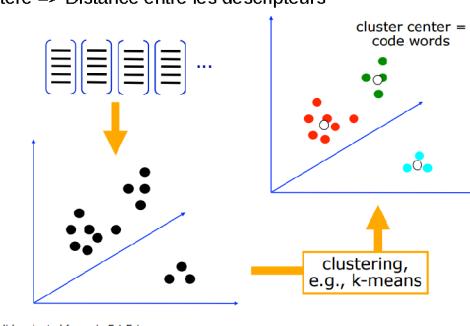


Cours IM3D – géométrie multi-vues - T59

ONERA
THE FRENCH AEROSPACE LAB

Apprentissage du dictionnaire

- Des descripteurs sont extraits à partir des images, qui sont classifiés à l'aide de méthodes statistiques.
- Critère => Distance entre les descripteurs



Cours IM3D – géométrie multi-vues - T57

ONERA
THE FRENCH AEROSPACE LAB

Résultats

Université d'Oxford

$p = 0.99$



Cours IM3D – géométrie multi-vues - T60

ONERA
THE FRENCH AEROSPACE LAB

Méthodes de SLAM visuel Dense et semi-dense

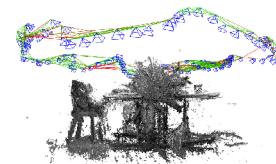
- Ne plus réaliser un appariement entre points d'intérêt mais entre des zones des l'images
- Calcul de pose qui résulte d'une optimisation sur un critère image de matching dense (cf cours appareillement dense)
- Plusieurs méthodes :
 - Capteurs : camera mono, capteur RGBD
 - Dense / Semi-dense
 - Différents critères d'association

Cours IM3D – IOGS 3A - T61

ONERA
THE FRENCH AEROSPACE LAB

LSD-SLAM

- Gestion de la carte
 - Critère basé sur la distances entre keyframes
 - Quand une nouvelle keyframe est créée, sa carte de profondeur est initialisée à en projetant la carte profondeur de la keyframe précédente.
 - A chaque itération, la keyframe est mise à jour à partir de l'image courante par calcul stéréo
- Détection de fermeture de boucle
 - Optimisation du graph de keyframe pour corriger la trajectoire globale

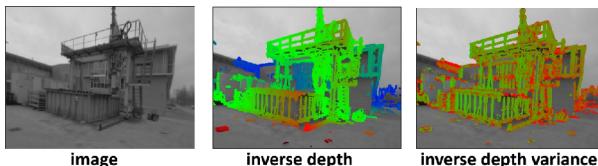


Cours IM3D – IOGS 3A - T64

ONERA
THE FRENCH AEROSPACE LAB

LSD-SLAM

- Algorithme de SLAM semi-dense mono-caméra
- Crée une carte composée de keyframes avec une structure de graphe
- Chaque keyframe est composée de l'image, de la carte d'inverse de la profondeur et de la carte de variance de l'inverse de la profondeur .
- Les cartes et sont définies uniquement pour un sous ensemble de pixels



LSD-SLAM: Large-Scale Direct Monocular SLAM

Cours IM3D – IOGS 3A - T62

ONERA

LSD-SLAM

- Résultats



Cours IM3D – IOGS 3A - T65

ONERA
THE FRENCH AEROSPACE LAB

LSD-SLAM

- Pour calculer la pose, on minimise le critère suivant :

$$E(\mathbf{R}, \mathbf{t}) = \sum_{x \in KF} (I_{KF}(x) - I(\omega(x, D_{KF}(x), \mathbf{R}, \mathbf{t}))^2$$
- Où ω est la fonction de warping entre l'image courante et la KF

$$\omega(p, d, \mathbf{R}, \mathbf{t}) = \begin{bmatrix} x'/z' \\ y'/z' \\ 1/z' \end{bmatrix} \text{ avec } \begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} p_x/d \\ p_y/d \\ 1/d \\ 1 \end{bmatrix}$$
- On remarque que ce critère est très similaire au critère de flot optique vue dans le cours d'appareillement dense.
- Optimisation du critère par Gauss-Newton pour déduire \mathbf{R} et \mathbf{t} .

Cours IM3D – IOGS 3A - T63

ONERA
THE FRENCH AEROSPACE LAB

Conclusion

- Méthodes d'odométrie visuelle permettent de reconstruire des trajectoires localement cohérentes
- Pour aller plus loin, les méthodes de SLAM permettent de créer une carte cohérente au niveau global
- Tendances
 - Les méthodes denses et semi-denses sont apparues récemment aidées par les progrès en informatique embarquée et commencent à concurrencer les approches points d'intérêts. Elles nécessitent des cadences caméras élevées (>30Hz)
 - Les méthodes de fermeture de boucles sont indispensables
 - Les approches par intégrations dériveront toujours
 - Vers une utilisation dans des applications industrielles ou destinées au grand public

Cours IM3D – IOGS 3A - T66

ONERA
THE FRENCH AEROSPACE LAB

Références

- **Odométrie visuelle**
 - D. Nistér, O. Naroditsky, James R. Bergen « Visual Odometry », proceedings of CVPR, 2004, pp. 560-567
 - M. Kaess, K. Ni, F. Dellaert, « Flow separation for fast and robust stereo odometry », proceedings of IEEE ICRA, 2009
 - D. Scaramuzza & F. Fraundorfer, « Visual Odometry: Part 1 – The first 30 years and fundamentals », IEEE Robotics and Automation Magazine, Vol. 18, Issue 4, 2011
 - F. Fraundorfer & D. Scaramuzza, « Visual Odometry: Part 2 – Matching, Robustness and applications », IEEE Robotics and Automation Magazine, Vol. 19, Issue 2, 2012
- **SLAM**
 - J. Civera, A.J. Davison, J.M. Martínez Montiel « Inverse Depth Parametrization for Monocular SLAM » in IEEE Transactions on robotics, Vol. 24, Issue 5, pp. 932-945, 2008
 - G. Klein, D. Murray « Parallel Tracking and Mapping for Small AR Workspaces» in proceedings of ISMAR, Nara, 2007
 - J. Engel, T. Schöps, D. Cremers, « LSD-SLAM: Large-Scale Direct Monocular SLAM » in ECCV, 2014
 - S. Thrun, W. Burgard, D. Fox « Probabilistic Robotics », The MIT Press, 2005