

Calibration spatiale et temporelle de systèmes multi-capteurs

I-Contexte et problématique

a- Introduction

Nous avons récemment travaillé sur des problématiques liées à la reconstruction 3D utilisant la technologie LiDAR. Le LiDAR est un capteur télémétrique qui donne en sortie un nuage de points représentant un sampling du monde physique qui l'entoure. Les coordonnées des points 3D donnés par le capteur LiDAR sont toujours exprimées dans un repère lié au LiDAR $\mathcal{R}_{lidar}(t)$.

$$\mathcal{R}_{lidar}(t) = (e_{xlidar}(t), e_{ylidar}(t), e_{zlidar}(t), T_{lidar}(t)) \quad (I.1)$$

On confondra par la suite la position du LiDAR avec l'origine de son repère de telle sorte que $T_{lidar}(t) = [T_x(t), T_y(t), T_z(t)]^T$ désignera la position du LiDAR exprimé dans un repère fixe.

Lorsque le capteur LiDAR se déplace dans le monde, il est possible d'avoir une reconstruction 3D de la scène en exprimant les coordonnées des points 3D fournis par le LiDAR dans un repère fixe \mathcal{R}_{world} .

$$\mathcal{R}_{world} = (e_x, e_y, e_z, O_{world}) \quad (I.2)$$

Pour ce faire, il nous faut connaître à tout instant la matrice de passage de la base B_{lidar} dans B_{world} :

$$P_{B_{lidar} \rightarrow B_{world}}(t) \quad (I.3)$$

Ainsi que la position du LiDAR. On alors, si X_{lidar} a été acquis à la date t :

$$X_{world} = P_{B_{lidar} \rightarrow B_{world}}(t) * X_{lidar} + (T_{lidar}(t) - O_{world}) \quad (I.4)$$

Dans le cas plus particulier où les bases B_{lidar} et B_{world} sont orthogonales direct, la matrice de passage de l'équation (I.3) prend une forme plus particulière.

En effet, le groupe spécial orthogonal de dimension 3 associé au produit scalaire canonique de \mathbb{R}^3 et restreint aux matrices dont le déterminant vaut 1 est constitué des matrices de rotations. En effet, si $u \in SO(\mathbb{R}^3)$ alors il existe une base orthonormée direct F telle que

$$Mat_F(u) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{pmatrix}$$

De mémoire, il suffit de montrer que $f : \lambda \rightarrow \det(u - \lambda Id)$ admet pour racine $\lambda_0 = 1$. Ainsi, il existe un vecteur fixe de u . On est donc ramené à $SO(\mathbb{R}^2)$ constitué des rotations.

On suppose à présent que les bases sont effectivement orthogonales direct.

On désigne par « orientation du LiDAR » la donnée de ses 3 angles d'Euler :

- Roll φ désigne la rotation autour de l'axe X : $\cos(\varphi) = \frac{\langle e_x; e_{xlidar} \rangle}{\|e_x\| \|e_{xlidar}\|}$, $\varphi \in [-\pi; \pi]$ (I.5)

- Pitch θ désigne la rotation autour de l'axe Y : $\cos(\theta) = \frac{\langle e_y; e_{ylidar} \rangle}{\|e_y\| \|e_{ylidar}\|}$, $\theta \in [-\frac{\pi}{2}; \frac{\pi}{2}]$ (I.6)

- Yaw ψ désigne la rotation autour de l'axe Z : $\cos(\psi) = \frac{\langle e_z; e_{zlidar} \rangle}{\|e_z\| \|e_{zlidar}\|}$, $\psi \in [-\pi; \pi]$ (I.7)

Si la position $T_{lidar}(t) = [T_x(t), T_y(t), T_z(t)]^T$ et l'orientation $\Theta_{lidar}(t) = [\varphi(t), \theta(t), \psi(t)]^T$ du lidar est connu à chaque instant alors il est possible d'exprimer directement les points générés par le lidar dans le repère \mathcal{R}_{world} :

On pose $R_{lidar \rightarrow world}(t) = R_{\psi(t)} * R_{\theta(t)} * R_{\varphi(t)}$

Avec,

$$R_{\psi(t)} = \begin{bmatrix} \cos(\psi(t)) & -\sin(\psi(t)) & 0 \\ \sin(\psi(t)) & \cos(\psi(t)) & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (I.8)$$

$$R_{\theta(t)} = \begin{bmatrix} \cos(\theta(t)) & 0 & \sin(\theta(t)) \\ 0 & 1 & 0 \\ -\sin(\theta(t)) & 0 & \cos(\theta(t)) \end{bmatrix} \quad (I.9)$$

$$R_{\varphi(t)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\varphi(t)) & -\sin(\varphi(t)) \\ 0 & \sin(\varphi(t)) & \cos(\varphi(t)) \end{bmatrix} \quad (I.10)$$

On a ainsi :

$$X_{world} = \begin{bmatrix} \frac{\|e_x\|}{\|e_{xlidar}\|} & 0 & 0 \\ 0 & \frac{\|e_y\|}{\|e_{ylidar}\|} & 0 \\ 0 & 0 & \frac{\|e_z\|}{\|e_{zlidar}\|} \end{bmatrix} * R_{lidar}(t) + (T_{lidar}(t) - O_{world}) \quad (I.11)$$

b-Contexte

Il nous faut à présent un moyen de connaître $\mathcal{R}_{lidar}(t)$ où une approximation suffisante à tout instant.

Pour cela on a recours à deux méthodes :

- Des instruments de mesures (centrale inertielle, GPS, ...)
- Des algorithmes d'odométrie comme le SLAM

Ces méthodes donnent des informations sur $\mathcal{R}_{lidar}(t)$ pour un ensemble de temps dénombrable fini.

Ainsi, on a accès aux informations :

$$\{\mathcal{R}_{lidar}(t_k)\}_{k \in [1, N]} \quad (I.11)$$

On approximera par la suite $\mathcal{R}_{lidar}(t)$ par des méthodes d'interpolation.

c-Problématique

Les clients avec lesquelles nous travaillons donnent en général les données suivantes :

- Nuages de points générés par le LiDAR
- Informations concernant le GPS
- Informations concernant la centrale inertielle

Et attendent de nous que nous leur fournissions l'ensemble des points de tous les nuages de points exprimés dans le repère \mathcal{R}_{world} .

Problème 1 : calibration spatiale

Le problème est que les clients oublient souvent de nous fournir les informations concernant les positions relatives des différents appareils de mesures. Cela entraîne des effets indésirables sur la reconstruction. Prenons l'exemple d'un avion embarquant un IMU et un GPS :

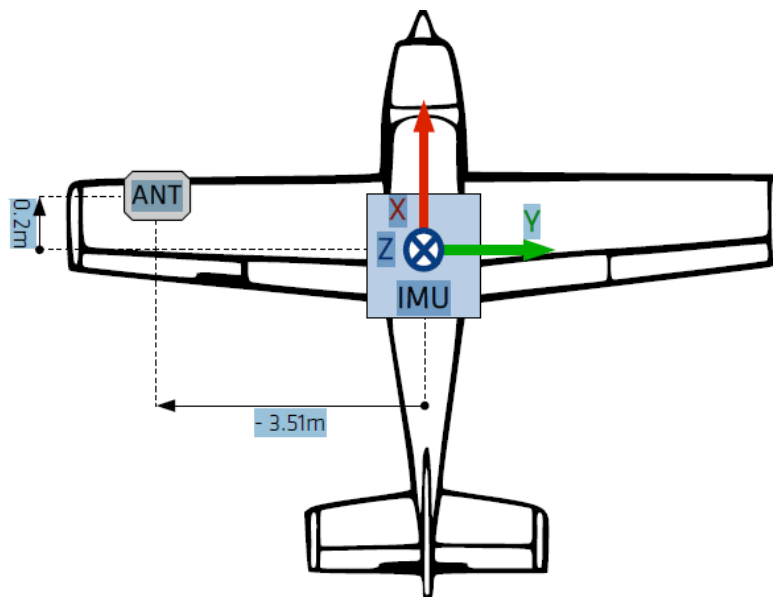


Fig I.1 : Exemple 1

Dans cet exemple, la position du GPS et celle de l'IMU ne sont pas confondus. Ainsi, si l'avion subit une rotation autour de l'axe Z avec pour centre de rotation la position de l'IMU, le GPS enregistrera le déplacement suivant :

$$\begin{matrix} X_{GPS} \\ y(t) \end{matrix} = \begin{matrix} x(t) \\ r * \sin(\theta(t) + \theta_0) \end{matrix} = \begin{matrix} r * \cos(\theta(t) + \theta_0) \\ r * \sin(\theta(t) + \theta_0) \end{matrix} + X_{IMU} \quad (I.12)$$

C'est ce que l'on appelle **l'effet bras de levier**

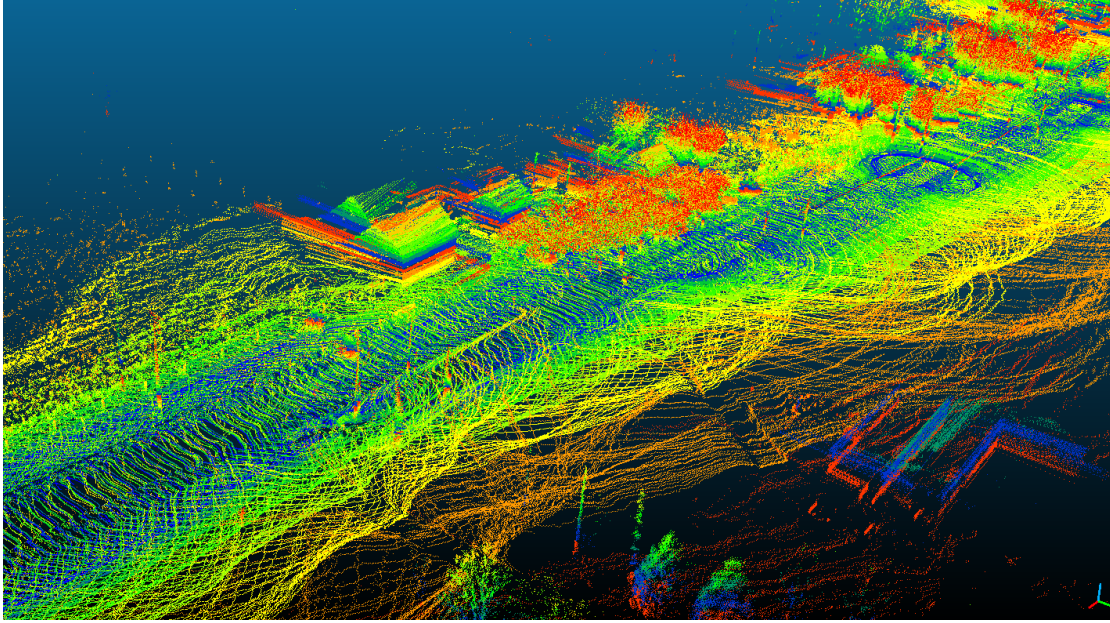


Fig 1.2 : Exemple d'effet bras de levier

La figure Fig1.2 montre l'effet d'un bras de levier sur une reconstruction. Ici, il s'agit d'un LiDAR, d'un GPS et d'un IMU montés sur le toit d'une voiture. La voiture traverse la rue une première fois, fait demi-tour puis retraverse la rue en sens inverse. Le GPS et le centre l'IMU n'étaient alignés ainsi pour une même position mais avec une rotation de 180° le GPS enregistre la translation suivante (l'axe X étant colinéaire à la direction de la route) :

$$X_{GPS}(\pi) - X_{GPS}(0) = \frac{r * (\cos(\pi) - \cos(0))}{r * (\sin(\pi) - \sin(0))} = \frac{0}{2r} \quad (1.13)$$

C'est pourquoi dans la figure 1.2 on observe une réplique de la scène avec un décalage égal à $2x$ la distance entre la position du GPS et de l'IMU.

Problème 2 : calibration temporelle

Bien souvent, les différents appareils de mesures utilise des timestamps différents. Ainsi on a :

$$t_{lidar} = s_{gps} * t_{gps} + t_0 = s_{IMU} * t_{IMU} + t_1 \quad (1.14)$$

Ce qui bien évidemment altère la qualité de la reconstruction (par exemple le véhicule tourne et la rotation est appliquée quelques secondes plus tard).

II- Pistes à explorer

a- Suppression de l'effet bras de levier

Problème 1

Supposons que nous ayons accès aux données suivantes :

- Position GPS du solide embarquant les différents capteurs (connaissance de X_{GPS}) dans un repère fixe de l'espace.
- Orientation du solide embarquant les différents capteurs (connaissance de Θ_{solide}) par rapport à un repère fixe de l'espace.
- Nuage de point généré par le LiDAR (coordonnées des points exprimés dans le repère \mathcal{R}_{lidar})

L'objectif est de retrouver les paramètres (r, θ_0, X_{IMU}) (voir équation I.12). On dispose pour cela des données suivantes :

$$\{X_{GPS}(t_k), \Theta_{IMU}(t_k)\}_{k \in [1, N]} \quad (II.1)$$

Cas bidimensionnelle :

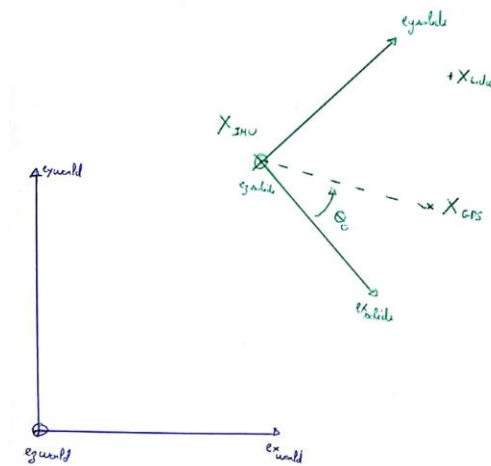


Fig II.1

Le problème est mal défini puisque qu'aucune solution n'existe. Pour s'en convaincre, imaginons les deux cas suivants :

- Le solide est immobile et subit une rotation pure
- Le solide est mobile et suit une trajectoire circulaire

Comment différencier dans ces deux cas l'effet bras de levier et celui d'un vrai mouvement circulaire ?

En l'état ce n'est pas possible.

Faisons alors l'hypothèse que toutes nouvelles positions du GPS se décompose de la façon suivante : une translation rectiligne plus une rotation due à l'effet bras de levier ?

On pose:

$$F(r, \theta_0) = \sum_{k=1}^N [r * \cos(\theta_k - \theta_0) + (x_{IMU} - x_{GPS})_k]^2 + [r * \sin(\theta_k - \theta_0) + (y_{IMU} - y_{GPS})_k]^2 \quad (II.2)$$

Problème 2

Supposons que nous ayons accès aux données suivantes :

- Position GPS du solide embarquant les différents capteurs (connaissance de X_{GPS}) dans un repère fixe de l'espace.
- Position de l'IMU du solide embarquant les différents capteurs (via un SLAM par exemple)
- Orientation du solide embarquant les différents capteurs (connaissance de Θ_{solide}) par rapport à un repère fixe de l'espace.
- Nuage de point généré par le LiDAR (coordonnées des points exprimés dans le repère \mathcal{R}_{lidar})

L'objectif est de retrouver les paramètres (r, θ_0) (voir équation I.12). On dispose pour cela des données suivantes :

$$\{X_{GPS}(t_k), X_{IMU}(t_k), \Theta_{IMU}(t_k)\}_{k \in [1, N]} \quad (II.1)$$

Les données $X_{GPS}(t_k)$ et $X_{IMU}(t_k)$ sont entachés d'erreurs. Pour retrouver les paramètres du problème on souhaite maximiser la vraisemblance :

$$P(X_{GPS}, X_{IMU} | r, \theta_0, \Theta_{IMU}) \quad (II.2)$$

b- Suppression du décalage temporel

Soit $L^2(\mathbb{R})$ l'ensemble des fonctions de carré intégrable au sens de la mesure de Lebesgue

Supposons que nous ayons accès aux données suivantes :

- Une fonction $f \in L^2(\mathbb{R})$ qui représente une mesure d'une dimension physique (position, orientation, ...)
- Une fonction $g \in L^2(\mathbb{R})$ qui représente la même mesure à l'aide d'un capteur / algorithme différent.

Par exemple, f peut-être la coordonnée X de la position du capteur dans un référentiel fixe fourni par un GPS et g la même mesure fourni par un algorithme de SLAM. En règle général, on se souciera peu de démontrer l'appartenance de f et g à $L^2(\mathbb{R})$ puisque ces fonctions sont à support fini, leurs carré est donc sommable.

Puisque nos capteurs ont des origines et échelles de temps différentes il vient que :

$$g(t) = f(s * t + \tau) + \varepsilon(t) \quad (II.3)$$

Où $\varepsilon(t)$ représente le bruit de la mesure.

Problème 1 : même échelle

On suppose dans un premier temps que le facteur d'échelle vaut 1 :

$$g(t) = f(t + \tau) + \varepsilon(t) \quad (II.3)$$

On souhaite estimer le paramètre τ . On définit τ^* :

$$\tau^* = \operatorname{argmin}_{\tau} \int \|f(t) - g(t - \tau)\|^2 dt \quad (II.4)$$

Avec $\|f\|$ la norme canonique de $L^2(\mathbb{R})$ associée au produit scalaire :

$$\langle f, g \rangle = \int f g \quad (11.5)$$

Ainsi, on a

$$\tau^* = \operatorname{argmin}_{\tau} \int f^2(t) + g(t - \tau)^2 - 2 * f(t) * g(t - \tau) dt$$

$$\tau^* = \operatorname{argmin}_{\tau} - 2 * \int f(t) * g(t - \tau) dt + \text{Const}$$

$$\tau^* = \operatorname{argmax}_{\tau} \int f(t) * g(t - \tau) dt$$

$$\tau^* = \operatorname{argmax}_{\tau} C_{f,g}(\tau) \quad (11.5)$$

Avec $C_{f,g}$ l'inter corrélation entre les fonctions f et g. Ainsi, il suffit de prendre le maximum de l'inter corrélation entre f et g afin de déterminer le décalage temporel entre ces deux signaux. Afin d'optimiser les temps de calculs ; l'inter corrélation sera calculé au moyen de la transformée de Fourier en remarquant que :

$$C_{f,g}(\tau) = \int f(t) * g(t - \tau) dt = \operatorname{conv}(f(t), g(-(t - \tau))) \quad (11.6)$$

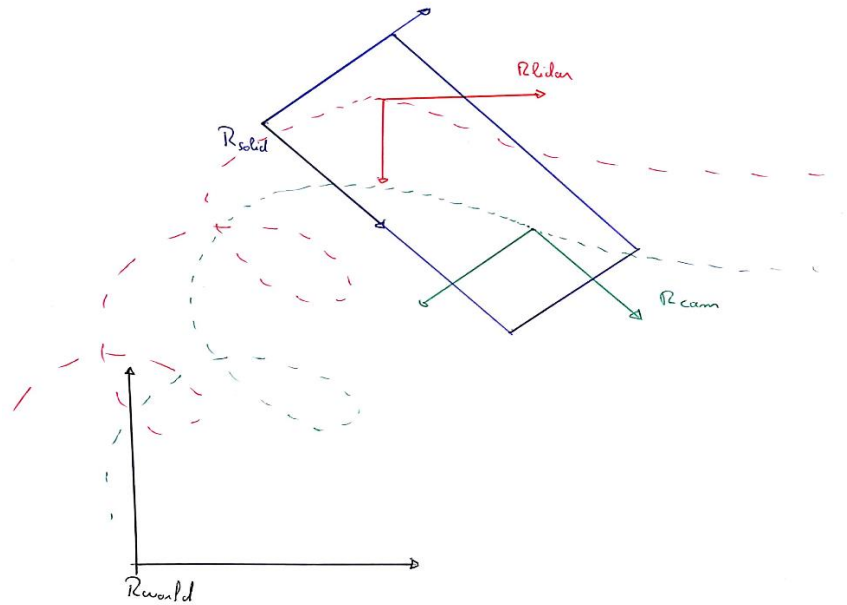
Ainsi :

$$\tilde{C}_{f,g}(\tau) = \tilde{F} * \widetilde{G(-)} = \tilde{F} * \widetilde{\operatorname{conj}(G)} \quad (11.7)$$

c- Calibration automatique Lidar-Caméra

Problème 1

On suppose ici que nous avons un lidar et une caméra fixés sur un solide de référence, le tout formant un solide (voir figure II.2)



II.1

Par la suite on notera :

$R_{ref1 \rightarrow ref2}$: la rotation permettant de passer de la base du référentiel 1 à celle du référentiel 2

$T_{ref1 \rightarrow ref2}$: la translation permettant de passer du référentiel 1 au référentiel 2

Dans un premier temps, on peut exprimer l'orientation et la trajectoire du lidar et de la caméra à partir de l'orientation et la trajectoire du solide de référence puisque le tout forme un solide :

$$X_{lidar}^{world}(t) = X_{solid}^{world}(t) + R_{solid}^{world}(t) * T_{lidar}^{solid}$$

$$R_{lidar}^{world} = R_{solid}^{world}(t) * R_{lidar}^{solid}$$

$$X_{cam}^{world}(t) = X_{solid}^{world}(t) + R_{solid}^{world}(t) * T_{cam}^{solid}$$

$$R_{cam}^{world} = R_{solid}^{world}(t) * R_{cam}^{solid}$$

NB 1 : T_{lidar}^{solid} et T_{cam}^{solid} ne dépendent pas du temps puisque le tout forme un solide

NB 2 : A chaque instant $X_{lidar}^{world}(t)$ est obtenu par une transformation affine à partir de $X_{solid}^{world}(t)$.

Cependant, la trajectoire $Traj_{lidar}^{world} = \{X_{lidar}^{world}(t)\}_{t \in I}$ n'est pas a priori une transformation affine de celle du solide puisque la rotation évolue au cours du temps.

Il est également intéressant de d'étudier la trajectoire du lidar dans le référentiel correspondant à son orientation et sa position à l'instant $t = 0$ puisque dans la pratique les trajectoires que nous auront seront issues des algorithmes de SLAM qui donnent une estimation de ces trajectoires.

$$X_{lidar}^{L0}(t) = R_{lidar}^{world^{-1}}(0) * (X_{lidar}^{world}(t) - X_{L0}^{world}(0))$$

$$R_{lidar}^{L0}(t) = R_{lidar}^{world^{-1}}(0) * R_{lidar}^{world}(t)$$

Voici à présent le problème auquel nous seront confronté :

Connaissant : $X_{lidar}^{L0}(t), R_{lidar}^{L0}(t), X_{cam}^{C0}(t)$ et $R_{lidar}^{C0}(t)$ on souhaite retrouver la position et orientation relative entre la caméra et le lidar : R_{L0}^{C0} et T_{L0}^{C0} .

Pour cela on va exprimer $X_{lidar}^{C0}(t)$ en effectuant un changement de référentiel d'une part ; et en utilisant le fait qu'ils forment un solide d'autre part :

$$X_{lidar}^{C0}(t) = X_{cam}^{C0}(t) + R_{cam}^{C0}(t) * T_{L0}^{C0} = R_{L0}^{C0} * X_{lidar}^{L0}(t) + T_{L0}^{C0}$$

C'est un système linéaire de la forme $AX = Y$ avec $X = [R_{L0}^{C0} | T_{L0}^{C0}]$