

# Propagations d’erreur pour l’ajustement de faisceaux local

## Error Propagations for Local Bundle Adjustment

A. Eudes<sup>1,2</sup>

M. Lhuillier<sup>1</sup>

<sup>1</sup> LASMEA UMR 6602, Université Blaise Pascal/CNRS , 63177 Aubière cedex, France

<sup>2</sup> LSVE, CEA/LIST/DTSI/SARC , 91191 Gif-sur-Yvette , France

Alexandre Eudes, LASMEA UMR 6602, 24 avenue des landais, 63177 Aubière cedex, France

Alexandre.eudes@lasmea.univ-bpclermont.fr

### Résumé

*L’ajustement de faisceaux local (AFL) a été récemment introduit pour estimer la géométrie à partir d’une séquence d’images prise par une caméra calibrée. Son avantage par rapport à la méthode standard d’ajustement de faisceaux (global) est une réduction importante de la complexité, ce qui permet des performances temps réel pour une précision similaire.*

*Cependant, aucune mesure de confiance sur les résultats de l’AFL comme une incertitude ou une covariance n’a été proposée jusqu’à présent. Cet article présente des modèles statistiques et des estimateurs pour le calcul d’incertitude en recherchant deux propriétés : (1) la propagation de l’incertitude tout au long de la séquence et (2) le calcul temps réel. Nous expliquons aussi pourquoi ce problème est plus compliqué qu’il n’y paraît, et nous donnons des résultats sur des données réelles.*

### Mots Clef

Ajustement de faisceau local, covariance, propagation d’erreur.

### Abstract

*Local bundle adjustment (LBA) has recently been introduced to estimate the geometry of image sequences taken by a calibrated camera. Its advantage over standard (global) bundle adjustment is a great reduction of computational complexity, which allows real-time performances with a similar accuracy. However, no confidence measure on the LBA result such as uncertainty or covariance has yet been introduced. This paper introduces statistical models and estimation methods for uncertainty with two desirable properties : (1) uncertainty propagation along the sequence and (2) real-time calculation. We also explain why this problem is more complicated than it may appear at first glance, and we provide results on video sequences.*

### Keywords

Local bundle adjustment, covariance, error propagation.

## 1 Introduction

L’estimation automatique et robuste du mouvement d’une caméra et des points de la scène à partir d’une séquence d’image (Structure-from-Motion ou **SfM**) est encore aujourd’hui un domaine de recherche actif.

Il y a trois ans, une méthode temps-réel basée sur un ajustement de faisceaux local (**AFL**) a été développée pour ce problème [10]. Cependant, aucune mesure de confiance sur le résultat n’a été donnée. Une telle mesure serait utile pour obtenir une information sur la qualité, si la vérité terrain n’est pas disponible, ou pour fusionner la reconstruction avec des données provenant d’autres capteurs comme un GPS ou un odomètre. Cette partie résume les travaux antérieurs, ainsi que notre contribution sur ce sujet.

**Ajustement de faisceaux** L’ajustement de faisceaux est une méthode bien connue [11] et conçue pour résoudre les problèmes de moindre carré non linéaire en SfM. On estime un vecteur  $\mathbf{x}$  qui minimise une fonction de coût  $\mathbf{x} \mapsto \|\mathbf{y} - F(\mathbf{x})\|^2$ . Le vecteur  $\mathbf{x}$  contient habituellement 6 paramètres pour chaque pose de caméra et 3 paramètres pour chaque point 3D. Le vecteur  $\mathbf{y}$  contient les primitives détectées dans les images,  $F$  est la concaténation des fonctions de projection, et  $\|\cdot\|$  est la norme Euclidienne. On appelle cette méthode ajustement de faisceaux global (**AFG**) si  $\mathbf{x}$  contient tous les paramètres de la séquence. La complexité d’une itération de l’AFG est  $O(c^3 + cp)$  avec  $c$  le nombre de poses de caméra et  $p$  le nombre de points 2D. Bien que cette complexité tienne compte de la structure creuse du problème en supposant  $c \ll p$ , un AFG temps réel est impossible sur de longues séquences.

**Propagation de l’erreur** la propagation d’erreur permet d’obtenir une mesure de la confiance sur les résultats de l’ajustement de faisceaux [8]. Supposons que le vecteur image  $\mathbf{y}$  suive un bruit gaussien connu. Alors, la fonction qui à  $\mathbf{y}$  associe le  $\mathbf{x}$  minimisant la fonction de coût est approximée par son développement de Taylor au premier ordre. Ainsi, Le bruit gaussien de  $\mathbf{x}$  est la propagation du bruit gaussien de  $\mathbf{y}$ . On peut donc estimer la matrice de covariance de  $\mathbf{x}$ . De cette façon, on définit les ellipses d’incertitude pour une probabilité donnée sur les résultats de

l'ajustement de faisceaux.

La complexité de l'estimation de la covariance pour toutes les poses de caméra et tous les points 3D est au moins celle d'une itération de l'AFG.

**Ajustement de faisceaux local** Nous avons récemment proposé l'AFL pour faire du SfM temps réel [10]. Cette méthode est incrémentale et est résumée de la façon suivante. Les poses de caméra et le nuage de points 3D sont estimés jusqu'à l'instant  $t - 1$ , et une nouvelle image doit être ajoutée à la reconstruction à l'instant  $t$ . Les points d'intérêts [7, 9] de l'image  $t$  sont détectés et appariés avec ceux de l'image  $t - 1$ . Comme plusieurs points de l'image  $t - 1$  sont reconstruits, il y a des correspondances 2D-3D pour les points de l'image  $t$  et la nouvelle pose peut être estimée.

Cette pose est initialisée par une méthode robuste [6], et l'AFL est appliqué pour améliorer la géométrie (poses de caméra et points 3D) des  $n$  dernières images, tout en maintenant la consistance grâce aux images précédentes  $t - N + 1, \dots, t - n$  avec  $n < N$ . L'AFL est un ajustement de faisceaux tel que (1)  $\mathbf{x}$  contient tout les paramètres 3D des  $n$  dernières images et (2)  $\mathbf{y}$  contient les primitives détectées dans les  $N$  dernières images.

La complexité d'une itération de l'AFL est  $O(n^3 + np)$  avec  $p$  le nombre de points 2D dans les  $N$  dernières images. La complexité de l'AFL est grandement réduite en utilisant des valeurs petites pour  $n$  et  $N$  (ex :  $n = 3, N = 10$ ).

**Notre contribution** Nous avons développé des modèles statistiques et des méthodes d'estimation de covariance pour les paramètres d'un système de SfM utilisant l'AFL. La partie 3 décrit la première méthode dérivée de la définition de l'AFL original utilisant une hypothèse d'indépendance forte (entre la pose des images précédentes  $t - N + 1, \dots, t - n$  et les primitives détectées dans les  $N$  dernières images). A noter que l'AFL original n'est pas un estimateur au maximum de vraisemblance (EMV) pour cette hypothèse. Ensuite, la Partie 4 décrit une seconde méthode dérivée d'un nouvel AFL qui est un EMV. Ces deux méthodes satisfont les propriétés désirées : (1) propagation de l'incertitude tout au long de la trajectoire et (2) des performances temps-réel. La partie 5 montre comment s'affranchir des hypothèses d'indépendance fortes supposées précédemment, mais les performances ne sont plus temps-réel. Enfin, nous montrerons des résultats sur des séquences vidéo et comparerons les covariances estimées avec celle provenant d'un AFG (partie 6).

## 2 Notations et définitions

**Temps et index d'image** La méthode SfM basée sur l'AFL distingue les images clefs des images non clef d'une séquence pour s'assurer de la stabilité de l'estimation 3D. Comme l'AFL est seulement utilisé sur les images clefs, on ignore les images non clefs et alors l'entier  $t$  est à la fois un index de temps et d'image clef.

**Poses de Caméras** Comme mentionné précédemment, l'AFL est un ajustement de faisceaux tel que les paramètres estimés  $\mathbf{x}$  sont tous les paramètres 3D des  $n$  dernières

images et le vecteur des mesures  $\mathbf{y}$  contient les points détectés dans les  $N$  dernières images. Pour simplifier les notations, sans perte de généralité, on utilisera  $n = 3$  et  $N = 10$ . Notez que la pose de la caméra définie pour une image est estimée à plusieurs reprises par plusieurs AFLs car la taille de la fenêtre glissante  $n$  est plus grande que un. Pour cette raison, un double indexage est utilisé pour les poses de caméra. Après l'AFL à l'itération  $t$ , la pose aux images  $0, 1, \dots, t$  est respectivement définie par le vecteur  $\mathbf{c}_t^0, \mathbf{c}_t^1, \dots, \mathbf{c}_t^t$ . Le vecteur  $\mathbf{c}_t^{t'}$  n'existe pas si  $t < t'$ . La méthode de SfM basée sur l'AFL procède en deux étapes :

-  $\exists t_0 > 0$  tel que  $\mathbf{c}_{t_0}^0, \mathbf{c}_{t_0}^1, \dots, \mathbf{c}_{t_0}^{t_0}$  est estimé par AFG.

-  $\forall t > t_0, \mathbf{c}_t^{t-2}, \mathbf{c}_t^{t-1}, \mathbf{c}_t^t$  est estimé par AFL à partir de  $\mathbf{c}_{t-1}^{t-9}, \mathbf{c}_{t-1}^{t-8}, \dots, \mathbf{c}_{t-1}^{t-3}$ .

La première étape est l'initialisation, elle fixe les coordonnées de toute la reconstruction. La seconde étape est incrémentale et ne modifie pas la pose aux images  $t - 9, t - 8, \dots, t - 3$ . En d'autres termes,  $\forall t' \leq t - 3, \mathbf{c}_t^{t'} = \mathbf{c}_{t-1}^{t'}$ .

**Ajustement de faisceau local** l'AFL au temps  $t$  estime aussi l'ensemble des points 3D  $\{\mathbf{s}_t^i\}$  qui ont au moins une projection dans les images  $t - 2, t - 1$  et  $t$ . Soit  $\mathbf{m}_t^{i'}$  la position détectée du point  $\mathbf{s}_t^i$  dans l'image  $t'$  tel que  $t' \in \{t - 9, t - 8, \dots, t\}$  (si elle existe). Les vecteurs  $\mathbf{c}_t^{t'}, \mathbf{s}_t^i, \mathbf{m}_t^{i'}$  ont pour dimension respective 6, 3 et 2. si  $\mathbf{x}$  et  $\mathbf{y}$  sont des vecteurs, on définit  $[\mathbf{xy}] = [\mathbf{x}|\mathbf{y}] = [\mathbf{x}^T \mathbf{y}^T]^T$  le vecteur qui concatène  $\mathbf{x}$  et  $\mathbf{y}$ . On définit

$$\mathbf{x}_t = [\mathbf{c}_t^{t-2} \mathbf{c}_t^{t-1} \mathbf{c}_t^t | \dots | \mathbf{s}_t^i \dots] \quad (1)$$

$$\mathbf{y}_t = [\dots | \mathbf{m}_t^{i'} \dots] \quad (2)$$

$$\mathbf{p}_t = [\mathbf{c}_t^{t-9} \mathbf{c}_t^{t-8} \dots \mathbf{c}_t^{t-3}] \quad (3)$$

ce qui permet d'écrire succinctement l'AFL :

$$\mathbf{x}_t = f_t(\mathbf{p}_t, \mathbf{y}_t) \text{ avec} \quad (4)$$

$$f_t(\mathbf{p}_t, \mathbf{y}_t) = \underset{\tilde{\mathbf{x}}_t}{\operatorname{argmin}} \|\mathbf{y}_t - F_t(\mathbf{p}_t, \tilde{\mathbf{x}}_t)\|^2. \quad (5)$$

La fonction  $F_t$  concatène les fonctions de projections des points 3D  $\{\mathbf{s}_t^i\}$  sur les points 2D  $\{\mathbf{m}_t^{i'}\}$  au bruit des images près. Différentes notations sont utilisées ici pour le minimum  $\mathbf{x}_t$  et le vecteur  $\tilde{\mathbf{x}}_t$  variable de la fonction  $F_t$ .

Ainsi,  $f_t$  minimise  $[\tilde{\mathbf{c}}_t^{t-2} \tilde{\mathbf{c}}_t^{t-1} \tilde{\mathbf{c}}_t^t | \sqrt{\tilde{\mathbf{s}}_t^i}] \mapsto$

$$\sum_{\forall i, 9 \leq t-t' \leq 3} \|\mathbf{m}_t^{i'} - \mathbf{c}_t^{t'} \tilde{\mathbf{s}}_t^i\|^2 + \sum_{\forall i, 2 \leq t-t' \leq 0} \|\mathbf{m}_t^{i'} - \tilde{\mathbf{c}}_t^{t'} \tilde{\mathbf{s}}_t^i\|^2 \quad (6)$$

avec  $\mathbf{c}_t^{t'} \tilde{\mathbf{s}}_t^i, \tilde{\mathbf{c}}_t^{t'} \tilde{\mathbf{s}}_t^i$  les projections de  $\tilde{\mathbf{s}}_t^i$  par les poses  $\mathbf{c}_t^{t'}$  et  $\tilde{\mathbf{c}}_t^{t'}$ .

**Autres Notations** Un sous-vecteur de  $\mathbf{x}$  est un vecteur obtenu en enlevant une ou plusieurs coordonnées de  $\mathbf{x}$ . Une sous-matrice d'une matrice carrée  $\mathbf{C}$  est une matrice obtenue à partir de  $\mathbf{C}$  en enlevant une ligne et une colonne de même indice, cette opération pouvant être réitérée pour plusieurs indices. On définit  $\mathbf{x}_t^c = [\mathbf{c}_t^{t-2} \mathbf{c}_t^{t-1} \mathbf{c}_t^t]$  comme un sous-vecteur de  $\mathbf{x}_t$ . La notation  $\mathbf{z} \sim \mathcal{N}(\bar{\mathbf{z}}, \mathbf{C}_z)$  signifie que  $\mathbf{z}$  est un vecteur gaussien de moyenne  $\bar{\mathbf{z}}$  et de covariance  $\mathbf{C}_z$ . Soit  $\mathbf{z} \mapsto g(\mathbf{z})$  une fonction  $\mathcal{C}^1$  continue de Jacobien  $\frac{\partial g}{\partial \mathbf{z}}$ . Au premier ordre près, on a la propagation d'erreur [8]

$$g(\mathbf{z}) \sim \mathcal{N}(g(\bar{\mathbf{z}}), \mathbf{C}_g) \text{ avec } \mathbf{C}_g = \frac{\partial g}{\partial \mathbf{z}}(\bar{\mathbf{z}}) \mathbf{C}_z \left( \frac{\partial g}{\partial \mathbf{z}}(\bar{\mathbf{z}}) \right)^T. \quad (7)$$

**Relation de récurrence** Maintenant, la méthode de SfM basée sur l'AFL peut s'écrire sous la forme d'une relation de récurrence sur les poses de caméra. On estime  $[\mathbf{p}_t \mathbf{x}_t^c]$  à partir de  $[\mathbf{p}_{t-1} \mathbf{x}_{t-1}^c]$  comme suit :

1.  $\mathbf{p}_t$  est un sous-vecteur de  $[\mathbf{p}_{t-1} \mathbf{x}_{t-1}^c]$
2. on estime  $\mathbf{x}_t = f_t(\mathbf{p}_t, \mathbf{y}_t)$  en utilisant l'AFL (Eq. 5)
3.  $\mathbf{x}_t^c$  est un sous-vecteur de  $\mathbf{x}_t$

L'étape 1 est évidente car  $[\mathbf{p}_{t-1} \mathbf{x}_{t-1}^c] = [\mathbf{c}_{t-1}^{t-10} \dots \mathbf{c}_{t-1}^{t-1}]$  et  $\mathbf{p}_t = [\mathbf{c}_t^{t-9} \dots \mathbf{c}_t^{t-3}] = [\mathbf{c}_{t-1}^{t-9} \dots \mathbf{c}_{t-1}^{t-3}]$ .

### 3 AFL original et sa covariance

Le but de cette partie est de faire le lien entre la covariance de  $[\mathbf{p}_t \mathbf{x}_t^c]$  et la relation de récurrence (cf partie 2) : on suppose que  $[\mathbf{p}_{t-1} \mathbf{x}_{t-1}^c] \sim \mathcal{N}([\bar{\mathbf{p}}_{t-1} \bar{\mathbf{x}}_{t-1}^c], \mathbf{C}_{[\mathbf{p}_{t-1} \mathbf{x}_{t-1}^c]})$  avec la covariance  $\mathbf{C}_{[\mathbf{p}_{t-1} \mathbf{x}_{t-1}^c]}$  connue et nous voulons estimer  $\mathbf{C}_{[\mathbf{p}_t \mathbf{x}_t^c]}$  tel que  $[\mathbf{p}_t \mathbf{x}_t^c] \sim \mathcal{N}([\bar{\mathbf{p}}_t \bar{\mathbf{x}}_t^c], \mathbf{C}_{[\mathbf{p}_t \mathbf{x}_t^c]})$ .

On suppose que  $\mathbf{y}_t \sim \mathcal{N}(\bar{\mathbf{y}}_t, \sigma^2 \mathbf{I})$  avec  $\sigma > 0$ , le niveau de bruit estimé par AFG durant l'initialisation de la méthode [8].

A première vue, la covariance  $\mathbf{C}_{\mathbf{x}_t}$  des paramètres  $\mathbf{x}_t$  estimés par AFL peut être approximée par l'inverse de l'approximation du hessien de la fonction de coût de l'AFL en  $\mathbf{x}_t$  [8]. Cependant, cette estimation ne propage pas le bruit depuis les paramètres  $\mathbf{p}_t$  précédents vers les nouveaux paramètres  $\mathbf{x}_t$ ; cela propage seulement le bruit de  $\mathbf{y}_t$  vers celui de  $\mathbf{x}_t$ . Malheureusement, cela implique que l'incertitude des poses de la caméra ne va pas croître avec le temps (ce n'est pas le résultat attendu). Notre méthode n'a pas ce problème, mais est plus compliquée.

#### 3.1 Modèle statistique de $[\mathbf{p}_t \mathbf{y}_t]$

La covariance  $\mathbf{C}_{\mathbf{p}_t}$  est une sous-matrice de  $\mathbf{C}_{[\mathbf{p}_{t-1} \mathbf{x}_{t-1}^c]}$  car  $\mathbf{p}_t$  est un sous-vecteur de  $[\mathbf{p}_{t-1} \mathbf{x}_{t-1}^c]$ . De plus, on a  $\mathbf{y}_t \sim \mathcal{N}(\bar{\mathbf{y}}_t, \sigma^2 \mathbf{I})$ . On suppose que les vecteurs gaussiens  $\mathbf{p}_t$  et  $\mathbf{y}_t$  sont indépendants et on obtient :

$$[\mathbf{p}_t \mathbf{y}_t] \sim \mathcal{N}([\bar{\mathbf{p}}_t \bar{\mathbf{y}}_t], \mathbf{C}_{[\mathbf{p}_t \mathbf{y}_t]}), \quad \mathbf{C}_{[\mathbf{p}_t \mathbf{y}_t]} = \begin{pmatrix} \mathbf{C}_{\mathbf{p}_t} & 0 \\ 0 & \sigma^2 \mathbf{I} \end{pmatrix}. \quad (8)$$

#### 3.2 Propagation de l'erreur

On approxime  $f_t$  par son développement de Taylor au premier ordre au point  $[\bar{\mathbf{p}}_t \bar{\mathbf{y}}_t]$  et on obtient

$$\begin{pmatrix} \mathbf{p} \\ f(\mathbf{p}, \mathbf{y}) \end{pmatrix} = \begin{pmatrix} \bar{\mathbf{p}} \\ f(\bar{\mathbf{p}}, \bar{\mathbf{y}}) \end{pmatrix} + \begin{pmatrix} \mathbf{I} & 0 \\ \frac{\partial f}{\partial \mathbf{p}} & \frac{\partial f}{\partial \mathbf{y}} \end{pmatrix} \begin{pmatrix} \mathbf{p} - \bar{\mathbf{p}} \\ \mathbf{y} - \bar{\mathbf{y}} \end{pmatrix}. \quad (9)$$

L'index  $t$  est omis dans cette expression. On en déduit que  $[\mathbf{p}_t \mathbf{x}_t] \sim \mathcal{N}([\bar{\mathbf{p}}_t \bar{\mathbf{x}}_t], \mathbf{C}_{[\mathbf{p}_t \mathbf{x}_t]})$  avec

$$\mathbf{C}_{[\mathbf{p}_t \mathbf{x}_t]} = \begin{pmatrix} \mathbf{I} & 0 \\ \frac{\partial f_t}{\partial \mathbf{p}_t} & \frac{\partial f_t}{\partial \mathbf{y}_t} \end{pmatrix} \mathbf{C}_{[\mathbf{p}_t \mathbf{y}_t]} \begin{pmatrix} \mathbf{I} & \frac{\partial f_t}{\partial \mathbf{p}_t} \\ 0 & \frac{\partial f_t}{\partial \mathbf{y}_t} \end{pmatrix}^T. \quad (10)$$

La proposition 1 explique comment estimer les dérivées de  $f_t$ .

**Proposition 1 :** Si  $F$  est  $\mathcal{C}^2$  et a une Jacobienne  $\frac{\partial F}{\partial \mathbf{x}}$  de rang plein et  $f(\mathbf{p}, \mathbf{y}) = \operatorname{argmin}_{\mathbf{x}} \|\mathbf{y} - F(\mathbf{p}, \mathbf{x})\|^2$ , alors

$$\frac{\partial f}{\partial \mathbf{y}} \approx \left( \frac{\partial F^T}{\partial \mathbf{x}} \frac{\partial F}{\partial \mathbf{x}} \right)^{-1} \frac{\partial F^T}{\partial \mathbf{x}} \quad (11)$$

$$\frac{\partial f}{\partial \mathbf{p}} \approx - \left( \frac{\partial F^T}{\partial \mathbf{x}} \frac{\partial F}{\partial \mathbf{x}} \right)^{-1} \frac{\partial F^T}{\partial \mathbf{x}} \frac{\partial F}{\partial \mathbf{p}} \quad (12)$$

avec les dérivées de  $F$  prises au point  $(\mathbf{p}, f(\mathbf{p}, \mathbf{y}))$ .

La preuve est donnée en Annexe A (l'index  $t$  est omis). Notons que l'Eq. 12 est une généralisation de l'équation bien connue Eq. 11. La matrice de covariance  $\mathbf{C}_{[\mathbf{p}_t \mathbf{x}_t^c]}$  est une sous-matrice en haut à gauche de  $\mathbf{C}_{[\mathbf{p}_t \mathbf{x}_t]}$ .

#### 3.3 Algorithme

On complète la relation de récurrence pour tenir compte des covariances. On estime  $[\mathbf{p}_t \mathbf{x}_t^c]$  et  $\mathbf{C}_{[\mathbf{p}_t \mathbf{x}_t^c]}$  à partir de  $[\mathbf{p}_{t-1} \mathbf{x}_{t-1}^c]$  et  $\mathbf{C}_{[\mathbf{p}_{t-1} \mathbf{x}_{t-1}^c]}$  comme suit :

1.  $\mathbf{p}_t$  est un sous-vecteur de  $[\mathbf{p}_{t-1} \mathbf{x}_{t-1}^c]$
2.  $\mathbf{C}_{\mathbf{p}_t}$  est une sous-matrice de  $\mathbf{C}_{[\mathbf{p}_{t-1} \mathbf{x}_{t-1}^c]}$
3. on estime  $\mathbf{x}_t = f_t(\mathbf{p}_t, \mathbf{y}_t)$  en utilisant l'AFL (Eq. 5)
4. on estime  $\mathbf{C}_{[\mathbf{p}_t \mathbf{x}_t]}$  en utilisant les Eqs. 8, 10, 11, et 12
5.  $\mathbf{C}_{[\mathbf{p}_t \mathbf{x}_t^c]}$  est un bloc haut-gauche de  $\mathbf{C}_{[\mathbf{p}_t \mathbf{x}_t]}$ .

Evidemment, on estime seulement le bloc  $\mathbf{C}_{[\mathbf{p}_t \mathbf{x}_t^c]}$  de  $\mathbf{C}_{[\mathbf{p}_t \mathbf{x}_t]}$  car les dimensions du vecteur caméra  $[\mathbf{p}_t \mathbf{x}_t^c]$  sont beaucoup plus petites que celle de  $[\mathbf{p}_t \mathbf{x}_t]$  qui contient aussi tout les paramètres des points 3D.

On donne maintenant les détails techniques de cette estimation. En combinant l'Eq. 10 et 8 on obtient

$$\mathbf{C}_{[\mathbf{p}_t \mathbf{x}_t]} = \begin{pmatrix} \mathbf{C}_{\mathbf{p}_t} & \mathbf{C}_{\mathbf{p}_t} \frac{\partial f_t}{\partial \mathbf{p}_t} \\ \frac{\partial f_t}{\partial \mathbf{p}_t} \mathbf{C}_{\mathbf{p}_t} & \frac{\partial f_t}{\partial \mathbf{p}_t} \mathbf{C}_{\mathbf{p}_t} \frac{\partial f_t}{\partial \mathbf{p}_t} + \sigma^2 \frac{\partial f_t}{\partial \mathbf{y}_t} \frac{\partial f_t}{\partial \mathbf{y}_t} \end{pmatrix} \quad (13)$$

Les dérivées de  $f_t$  sont données par l'Eq. 11 et 12. Elles sont estimées en  $(\mathbf{p}_t, \mathbf{x}_t)$  avec comme approximation  $(\mathbf{p}_t, \mathbf{x}_t) \approx (\bar{\mathbf{p}}_t, \bar{\mathbf{x}}_t)$ . Soit  $\mathbf{H}_t = \frac{\partial F_t^T}{\partial \mathbf{x}_t} \frac{\partial F_t}{\partial \mathbf{x}_t}$  l'approximation de Gauss-Newton du Hessian de la fonction de coût de l'AFL. Alors  $\mathbf{C}_{[\mathbf{p}_t \mathbf{x}_t^c]}$  est un bloc haut-gauche de  $\mathbf{C}_{[\mathbf{p}_t \mathbf{x}_t]}$  tel que

$$\mathbf{C}_{[\mathbf{p}_t \mathbf{x}_t^c]} = \begin{pmatrix} \mathbf{C}_{\mathbf{p}_t} & \mathbf{A}^T \\ \mathbf{A} & \mathbf{B} \end{pmatrix}, \quad \mathbf{B} \text{ est un bloc haut-gauche}$$

$$\text{de } \sigma^2 \mathbf{H}_t^{-1} + \mathbf{H}_t^{-1} \frac{\partial F_t^T}{\partial \mathbf{x}_t} \frac{\partial F_t}{\partial \mathbf{p}_t} \mathbf{C}_{\mathbf{p}_t} \frac{\partial F_t^T}{\partial \mathbf{p}_t} \frac{\partial F_t}{\partial \mathbf{x}_t} \mathbf{H}_t^{-1}$$

$$\text{et } \mathbf{A} \text{ est un bloc haut de } -\mathbf{H}_t^{-1} \frac{\partial F_t^T}{\partial \mathbf{x}_t} \frac{\partial F_t}{\partial \mathbf{p}_t} \mathbf{C}_{\mathbf{p}_t}.$$

Ensuite  $\mathbf{A}$  et  $\mathbf{B}$  sont estimées grâce à une inversion par bloc [11]

$$\mathbf{H}_t = \begin{pmatrix} \mathbf{U} & \mathbf{W} \\ \mathbf{W}^T & \mathbf{V} \end{pmatrix} \Rightarrow \mathbf{H}_t^{-1} = \begin{pmatrix} 0 & 0 \\ 0 & \mathbf{V}^{-1} \end{pmatrix} + \mathbf{Y}^T \mathbf{Z}^{-1} \mathbf{Y} \quad (14)$$

$$\text{avec } \mathbf{Z} = \mathbf{U} - \mathbf{W} \mathbf{V}^{-1} \mathbf{W}^T \text{ et } \mathbf{Y} = \begin{pmatrix} \mathbf{I} & -\mathbf{W} \mathbf{V}^{-1} \end{pmatrix}. \quad (15)$$

Les matrices  $B, U$  et  $Z$  ont les même dimensions ;  $V$  est une matrice diagonale par bloc avec des blocs  $3 \times 3$  inversible grâce à l'Eq. 6. On estime successivement  $Z^{-1}$ ,  $D_0 = \frac{\partial F_t^T}{\partial \mathbf{x}_t} \frac{\partial F_t}{\partial \mathbf{p}_t}$ , et  $D_1 = H_t^{-1} D_0$  en utilisant

$$D_1 = \begin{pmatrix} 0 & 0 \\ 0 & V^{-1} \end{pmatrix} D_0 + (Y^T (Z^{-1} (Y D_0))). \quad (16)$$

Ensuite, l'estimation de  $A$  et  $B$  est aisée : la matrice  $A$  est un bloc haut de  $-D_1 C_{\mathbf{p}_t}$  et  $B = \sigma^2 Z^{-1}$  et un bloc haut-gauche de  $D_1 C_{\mathbf{p}_t} D_1^T$ .

## 4 Nouvel AFL et sa covariance

Dans la partie 3, la covariance de l'AFL original [10] est estimée, en supposant que  $\mathbf{p}_t$  et  $\mathbf{y}_t$  sont indépendants. Dans la partie 4, on introduit une nouvelle fonction AFL  $f_t^n$  qui permet une estimation au maximum de vraisemblance (EMV) de  $\mathbf{x}_t$  sous ces hypothèses. Dans cette même partie, on donne la relation de récurrence permettant d'estimer la géométrie et sa covariance :  $[\mathbf{p}_t \mathbf{x}_t^c]$  et  $C_{[\mathbf{p}_t \mathbf{x}_t^c]}$  sont estimés à partir de  $[\mathbf{p}_{t-1} \mathbf{x}_{t-1}^c]$  et  $C_{[\mathbf{p}_{t-1} \mathbf{x}_{t-1}^c]}$ .

### 4.1 Estimation au maximum de vraisemblance

L'index  $t$  est omis pour  $\mathbf{p}_t, \tilde{\mathbf{p}}_t, \mathbf{x}_t, \tilde{\mathbf{x}}_t, \mathbf{y}_t, F_t, f_t$ . On suppose que  $\mathbf{p} \sim \mathcal{N}(\bar{\mathbf{p}}, C_{\mathbf{p}})$  et  $\mathbf{y} \sim \mathcal{N}(\bar{\mathbf{y}}, \sigma^2 \mathbf{I})$  sont indépendants avec  $\bar{\mathbf{y}} = F(\bar{\mathbf{p}}, \bar{\mathbf{x}})$ . Les paramètres inconnus du modèle statistique sont  $\bar{\mathbf{p}}$  et  $\bar{\mathbf{x}}$  (on suppose ici que  $\sigma^2$  et  $C_{\mathbf{p}}$  sont connus et non bruités). La fonction densité de probabilité du vecteur gaussien  $[\mathbf{p} \mathbf{y}]$  est

$$d(\mathbf{p}, \mathbf{y} | \bar{\mathbf{p}}, \bar{\mathbf{x}}) = K e^{-\frac{1}{2\sigma^2} (\|\mathbf{y} - F(\bar{\mathbf{p}}, \bar{\mathbf{x}})\|^2 + (\mathbf{p} - \bar{\mathbf{p}})^T C_{\mathbf{p}}^{-1} (\mathbf{p} - \bar{\mathbf{p}}))} \quad (17)$$

avec  $K$  une constante. Alors, l'EMV  $f_t^n(\mathbf{p}, \mathbf{y})$  de  $[\bar{\mathbf{p}} \bar{\mathbf{x}}]$  est

$$\operatorname{argmin}_{[\bar{\mathbf{p}} \bar{\mathbf{x}}]} \frac{1}{\sigma^2} \|\mathbf{y} - F(\bar{\mathbf{p}}, \bar{\mathbf{x}})\|^2 + (\mathbf{p} - \bar{\mathbf{p}})^T C_{\mathbf{p}}^{-1} (\mathbf{p} - \bar{\mathbf{p}}). \quad (18)$$

La fonction  $f_t^n$  définit les estimations  $\mathbf{p}_t^n$  et  $\mathbf{x}_t^n$  de  $\bar{\mathbf{p}}_t$  et  $\bar{\mathbf{x}}_t$  (notons que  $f_t$  ne donne pas d'estimation de  $\bar{\mathbf{p}}_t$ ).

### 4.2 Propagation de l'erreur

La fonction  $f_t^n(\mathbf{p}_t, \mathbf{y}_t)$  dans l'Eq. 18 peut se réécrire comme

$$f_t^n(\mathbf{y}_t^n) = \operatorname{argmin}_{\tilde{\mathbf{x}}_t^n} \|\mathbf{y}_t^n - F_t^n(\tilde{\mathbf{x}}_t^n)\|^2 \quad (19)$$

$$\text{avec } \tilde{\mathbf{x}}_t^n = [\tilde{\mathbf{p}}_t | \tilde{\mathbf{x}}_t], \quad \mathbf{y}_t^n = [(C_{\mathbf{p}_t})^{-\frac{1}{2}} \mathbf{p}_t | \frac{1}{\sigma} \mathbf{y}_t] \quad (20)$$

$$\text{et } F_t^n(\tilde{\mathbf{x}}_t^n) = [(C_{\mathbf{p}_t})^{-\frac{1}{2}} \tilde{\mathbf{p}}_t | \frac{1}{\sigma} F_t(\tilde{\mathbf{p}}_t, \tilde{\mathbf{x}}_t)]. \quad (21)$$

On voit maintenant que la jacobienne  $\frac{\partial f_t^n}{\partial \mathbf{y}_t^n}$  de  $f_t^n$  peut être estimée en utilisant l'Eq. 11 de la Proposition 1 :

$$\frac{\partial f_t^n}{\partial \mathbf{y}_t^n} = \left( \frac{\partial F_t^n}{\partial \tilde{\mathbf{x}}_t^n} \frac{\partial F_t^n}{\partial \tilde{\mathbf{p}}_t^n} \right)^{-1} \frac{\partial F_t^n}{\partial \tilde{\mathbf{x}}_t^n} \quad (22)$$

De plus, l'Eq. 20 implique  $\mathbf{y}_t^n \sim \mathcal{N}(\bar{\mathbf{y}}_t^n, \mathbf{I})$ . L'Eq. 22 et le développement de Taylor au premier ordre de  $f_t^n$  en  $[\tilde{\mathbf{p}}_t \tilde{\mathbf{y}}_t]$  donne

$$C_{[\mathbf{p}_t^n \mathbf{x}_t^n]} = C_{f_t^n} = \frac{\partial f_t^n}{\partial \mathbf{y}_t^n} \mathbf{I} \frac{\partial f_t^n}{\partial \mathbf{y}_t^n}^T = \left( \frac{\partial F_t^n}{\partial \tilde{\mathbf{x}}_t^n} \frac{\partial F_t^n}{\partial \tilde{\mathbf{p}}_t^n} \right)^{-1}. \quad (23)$$

### 4.3 Algorithme

On peut estimer  $[\mathbf{p}_t \mathbf{x}_t^c]$  et  $C_{[\mathbf{p}_t \mathbf{x}_t^c]}$  depuis  $[\mathbf{p}_{t-1} \mathbf{x}_{t-1}^c]$  et  $C_{[\mathbf{p}_{t-1} \mathbf{x}_{t-1}^c]}$  comme suit :

1.  $\mathbf{p}_t$  est un sous-vecteur de  $[\mathbf{p}_{t-1} \mathbf{x}_{t-1}^c]$
2.  $C_{\mathbf{p}_t}$  est une sous-matrice de  $C_{[\mathbf{p}_{t-1} \mathbf{x}_{t-1}^c]}$
3. on estime  $[\mathbf{p}_t^n \mathbf{x}_t^n] = f_t^n(\mathbf{p}_t, \mathbf{y}_t)$  en utilisant l'AFL (Eq. 18)
4. on estime  $C_{[\mathbf{p}_t^n \mathbf{x}_t^n]}$  en utilisant les Eqs. 23 et 21
5. on fait  $\mathbf{p}_t \leftarrow \mathbf{p}_t^n, \mathbf{x}_t \leftarrow \mathbf{x}_t^n$  et  $C_{[\mathbf{p}_t \mathbf{x}_t]} \leftarrow C_{[\mathbf{p}_t^n \mathbf{x}_t^n]}$
6.  $C_{[\mathbf{p}_t \mathbf{x}_t^c]}$  est un bloc haut-gauche de  $C_{[\mathbf{p}_t \mathbf{x}_t]}$ .

Le nouvel AFL (Eq. 18) est plus coûteux en temps de calcul que l'AFL original (Eq. 5) car les poses de caméras de  $\mathbf{p}_t^n$  doivent être estimées avec celles de  $\mathbf{x}_t^n$ . Donc on remplace l'étape 3 par  $\mathbf{x}_t = f_t(\mathbf{p}_t, \mathbf{y}_t)$  et on supprime l'étape 5. L'étape 4 reste inchangée : Eqs. 23 et 21 sont toujours utilisées avec  $\mathbf{p}_t^n = \mathbf{p}_t, \mathbf{x}_t^n = \mathbf{x}_t, C_{[\mathbf{p}_t \mathbf{x}_t]} = C_{[\mathbf{p}_t^n \mathbf{x}_t^n]}$ . En d'autres termes, on approxime le résultat du nouvel AFL par le résultat de l'AFL original (les deux résultats sont exactement les mêmes si il n'y a pas de bruit dans les images). Evidemment, on estime seulement le bloc  $C_{[\mathbf{p}_t \mathbf{x}_t^c]}$  de  $C_{[\mathbf{p}_t \mathbf{x}_t]}$  car la dimension du vecteur caméra  $[\mathbf{p}_t \mathbf{x}_t^c]$  est beaucoup plus petite que celle de  $[\mathbf{p}_t \mathbf{x}_t]$  qui contient aussi les paramètres des points 3D.

Maintenant, on donne les détails techniques de cette estimation. Les Eqs. 21 et 23 donnent ( $t$  et  $\sim$  omis)

$$\frac{\partial F^n}{\partial \mathbf{x}^n} = \begin{pmatrix} \frac{\partial F^n}{\partial \mathbf{p}} & \frac{\partial F^n}{\partial \mathbf{x}} \end{pmatrix} = \begin{pmatrix} C_{\mathbf{p}}^{-\frac{1}{2}} & 0 \\ \frac{1}{\sigma} \frac{\partial F}{\partial \mathbf{p}} & \frac{1}{\sigma} \frac{\partial F}{\partial \mathbf{x}} \end{pmatrix} \quad (24)$$

$$C_{[\mathbf{p}_t^n \mathbf{x}_t^n]} = \begin{pmatrix} C_{\mathbf{p}}^{-1} + \frac{1}{\sigma^2} \frac{\partial F^T}{\partial \mathbf{p}} \frac{\partial F}{\partial \mathbf{p}} & \frac{1}{\sigma^2} \frac{\partial F^T}{\partial \mathbf{p}} \frac{\partial F}{\partial \mathbf{x}} \\ \frac{1}{\sigma^2} \frac{\partial F^T}{\partial \mathbf{x}} \frac{\partial F}{\partial \mathbf{p}} & \frac{1}{\sigma^2} \frac{\partial F^T}{\partial \mathbf{x}} \frac{\partial F}{\partial \mathbf{x}} \end{pmatrix}^{-1} \quad (25)$$

On applique l'inversion par bloc (Eqs. 14 et 15) à  $C_{[\mathbf{p}_t^n \mathbf{x}_t^n]}$  avec  $\mathbf{x} = (\mathbf{x}^c \quad \mathbf{x}^s)$ ,

$$U = \begin{pmatrix} C_{\mathbf{p}}^{-1} + \frac{1}{\sigma^2} \frac{\partial F^T}{\partial \mathbf{p}} \frac{\partial F}{\partial \mathbf{p}} & \frac{1}{\sigma^2} \frac{\partial F^T}{\partial \mathbf{p}} \frac{\partial F}{\partial \mathbf{x}^c} \\ \frac{1}{\sigma^2} \frac{\partial F^T}{\partial \mathbf{x}^c} \frac{\partial F}{\partial \mathbf{p}} & \frac{1}{\sigma^2} \frac{\partial F^T}{\partial \mathbf{x}^c} \frac{\partial F}{\partial \mathbf{x}^c} \end{pmatrix} \quad (26)$$

$$W = \frac{1}{\sigma^2} \begin{pmatrix} \frac{\partial F}{\partial \mathbf{p}} & \frac{\partial F}{\partial \mathbf{x}^c} \end{pmatrix}^T \frac{\partial F}{\partial \mathbf{x}^s}, \quad V = \frac{1}{\sigma^2} \frac{\partial F^T}{\partial \mathbf{x}^s} \frac{\partial F}{\partial \mathbf{x}^s} \quad (27)$$

et on obtient  $C_{[\mathbf{p}_t \mathbf{x}_t^c]} = (U - WV^{-1}W^T)^{-1}$ . Grâce à l'Eq. 6,  $V$  est une matrice diagonal par bloc ( $3 \times 3$ ) et est facilement inversible.

## 5 Hypothèse faible

Les parties précédentes 3 et 4 nécessitent que  $\mathbf{p}_t \sim \mathcal{N}(\bar{\mathbf{p}}_t, \mathbf{C}_{\mathbf{p}_t})$  et  $\mathbf{y}_t \sim \mathcal{N}(\bar{\mathbf{y}}_t, \sigma^2 \mathbf{I})$  soient indépendants. Cependant, la pose  $\mathbf{c}_t^{t-3} = \mathbf{c}_{t-1}^{t-3}$  est à la fois sous-vecteur de  $\mathbf{p}_t$  et  $\mathbf{x}_{t-1}$  tel que  $\mathbf{x}_{t-1} = f_{t-1}(\mathbf{p}_{t-1}, \mathbf{y}_{t-1})$ . Comme  $\mathbf{y}_t$  et  $\mathbf{y}_{t-1}$  ont des points 2D en commun dans les images  $t-9, t-8, \dots, t-1$ , on ne peut pas supposer que  $\mathbf{p}_t$  et  $\mathbf{y}_t$  sont indépendants.

Dans cette partie 5, des relations de récurrence sont introduites pour estimer la covariance de l'AFL original et du nouvel AFL sans cette hypothèse :  $[\mathbf{y}_t \mathbf{p}_t \mathbf{x}_t^c]$  et  $\mathbf{C}_{[\mathbf{y}_t \mathbf{p}_t \mathbf{x}_t^c]}$  sont estimés à partir de  $[\mathbf{y}_{t-1} \mathbf{p}_{t-1} \mathbf{x}_{t-1}^c]$  et  $\mathbf{C}_{[\mathbf{y}_{t-1} \mathbf{p}_{t-1} \mathbf{x}_{t-1}^c]}$ .

### 5.1 Modèle statistique de $[\mathbf{y}_t \mathbf{p}_t]$

De nouvelles notations sont nécessaires ici. Soit  $\mathbf{y}_t \cap \mathbf{y}_{t-1}$  (respectivement,  $\mathbf{y}_t \setminus \mathbf{y}_{t-1}$ ) le sous-vecteur de  $\mathbf{y}_t$  avec les points 2D qui sont (respectivement, qui ne sont pas) dans  $\mathbf{y}_{t-1}$ . La covariance  $\mathbf{C}_{[\mathbf{y}_{t-1} \mathbf{p}_{t-1} \mathbf{x}_{t-1}^c]}$  est connue, et la covariance  $\mathbf{C}_{[\mathbf{y}_t \cap \mathbf{y}_{t-1} | \mathbf{p}_t]}$  est une sous-matrice de  $\mathbf{C}_{[\mathbf{y}_{t-1} \mathbf{p}_{t-1} \mathbf{x}_{t-1}^c]}$  puisque  $\mathbf{y}_t \cap \mathbf{y}_{t-1}$  est un sous-vecteur de  $\mathbf{y}_{t-1}$  et  $\mathbf{p}_t$  est un sous-vecteur de  $[\mathbf{p}_{t-1} \mathbf{x}_{t-1}^c]$ .

On suppose que  $\mathbf{y}_t \setminus \mathbf{y}_{t-1}$  et  $[\mathbf{y}_t \cap \mathbf{y}_{t-1} | \mathbf{p}_t]$  sont indépendants. Donc  $[\mathbf{y}_t \mathbf{p}_t]$  est un vecteur gaussien de covariance

$$\mathbf{C}_{[\mathbf{y}_t \mathbf{p}_t]} = \mathbf{C}_{[\mathbf{y}_t \setminus \mathbf{y}_{t-1} | \mathbf{y}_t \cap \mathbf{y}_{t-1} | \mathbf{p}_t]} = \begin{pmatrix} \sigma^2 \mathbf{I} & 0 \\ 0 & \mathbf{C}_{[\mathbf{y}_t \cap \mathbf{y}_{t-1} | \mathbf{p}_t]} \end{pmatrix}. \quad (28)$$

### 5.2 Propagation de l'erreur (AFL original)

On approxime  $f_t$  par son développement de Taylor au premier ordre au point  $[\bar{\mathbf{y}}_t \bar{\mathbf{p}}_t]$  et on obtient

$$\begin{pmatrix} \mathbf{y} \\ \mathbf{p} \\ f(\mathbf{p}, \mathbf{y}) \end{pmatrix} = \begin{pmatrix} \bar{\mathbf{y}} \\ \bar{\mathbf{p}} \\ f(\bar{\mathbf{p}}, \bar{\mathbf{y}}) \end{pmatrix} + \begin{pmatrix} \mathbf{I} & 0 \\ 0 & \mathbf{I} \\ \frac{\partial f}{\partial \mathbf{y}} & \frac{\partial f}{\partial \mathbf{p}} \end{pmatrix} \begin{pmatrix} \mathbf{y} - \bar{\mathbf{y}} \\ \mathbf{p} - \bar{\mathbf{p}} \end{pmatrix}. \quad (29)$$

L'index  $t$  est omis dans cette expression. On déduit que  $[\mathbf{y}_t \mathbf{p}_t \mathbf{x}_t] \sim \mathcal{N}([\bar{\mathbf{y}}_t \bar{\mathbf{p}}_t \bar{\mathbf{x}}_t], \mathbf{C}_{[\mathbf{y}_t \mathbf{p}_t \mathbf{x}_t]})$  avec sa covariance

$$\mathbf{C}_{[\mathbf{y}_t \mathbf{p}_t \mathbf{x}_t]} = \begin{pmatrix} \mathbf{I} & 0 \\ 0 & \mathbf{I} \\ \frac{\partial f_t}{\partial \mathbf{y}_t} & \frac{\partial f_t}{\partial \mathbf{p}_t} \end{pmatrix} \mathbf{C}_{[\mathbf{y}_t \mathbf{p}_t]} \begin{pmatrix} \mathbf{I} & 0 \\ 0 & \mathbf{I} \\ \frac{\partial f_t^T}{\partial \mathbf{y}_t} & \frac{\partial f_t^T}{\partial \mathbf{p}_t} \end{pmatrix}. \quad (30)$$

Les dérivées de  $f_t$  sont données par la Proposition 1.

### 5.3 Algorithme (AFL original)

Les parties 5.2 et 5.1 définissent la relation de récurrence pour l'AFL original. On estime  $[\mathbf{y}_t \mathbf{p}_t \mathbf{x}_t^c]$  et  $\mathbf{C}_{[\mathbf{y}_t \mathbf{p}_t \mathbf{x}_t^c]}$  grâce à  $[\mathbf{y}_{t-1} \mathbf{p}_{t-1} \mathbf{x}_{t-1}^c]$  et  $\mathbf{C}_{[\mathbf{y}_{t-1} \mathbf{p}_{t-1} \mathbf{x}_{t-1}^c]}$  comme suit :

1.  $\mathbf{p}_t$  est un sous-vecteur de  $[\mathbf{y}_{t-1} \mathbf{p}_{t-1} \mathbf{x}_{t-1}^c]$
2.  $\mathbf{C}_{[\mathbf{y}_t \cap \mathbf{y}_{t-1} | \mathbf{p}_t]}$  est une sous-matrice de  $\mathbf{C}_{[\mathbf{y}_{t-1} \mathbf{p}_{t-1} \mathbf{x}_{t-1}^c]}$
3. on estime  $\mathbf{x}_t = f_t(\mathbf{p}_t, \mathbf{y}_t)$  en utilisant l'AFL (Eq. 5)
4. on estime  $\mathbf{C}_{[\mathbf{y}_t \mathbf{p}_t \mathbf{x}_t]}$  en utilisant Eqs. 28, 30, 11, 12
5.  $\mathbf{C}_{[\mathbf{y}_t \mathbf{p}_t \mathbf{x}_t^c]}$  est un bloc haut-gauche de  $\mathbf{C}_{[\mathbf{y}_t \mathbf{p}_t \mathbf{x}_t]}$ .

Cette fois-ci, l'estimation complète de  $\frac{\partial f_t}{\partial \mathbf{y}_t}$  par l'Eq. 11 est requise pour le bloc haut-gauche  $\mathbf{C}_{[\mathbf{y}_t \mathbf{p}_t \mathbf{x}_t^c]}$  de  $\mathbf{C}_{[\mathbf{y}_t \mathbf{p}_t \mathbf{x}_t]}$  et la méthode devient inutilisable en temps-réel.

## 5.4 Estimateur au maximum de vraisemblance (nouvel AFL)

L'index  $t$  est omis pour  $\mathbf{p}_t, \bar{\mathbf{p}}_t, \mathbf{x}_t, \bar{\mathbf{x}}_t, \mathbf{y}_t, F_t, f_t$  dans cette partie. On suppose que  $[\mathbf{y} \mathbf{p}] \sim \mathcal{N}([\bar{\mathbf{y}} \bar{\mathbf{p}}], \mathbf{C}_{[\mathbf{y} \mathbf{p}]})$  avec  $\mathbf{C}_{[\mathbf{y} \mathbf{p}]}$  définie par l'Eq. 28 et  $\bar{\mathbf{y}} = F(\bar{\mathbf{p}}, \bar{\mathbf{x}})$ . Les paramètres inconnus du modèle statistique sont  $\bar{\mathbf{p}}$  et  $\bar{\mathbf{x}}$  (on suppose que  $\mathbf{C}_{[\mathbf{y} \mathbf{p}]}$  est connue et non bruité). La fonction densité de probabilité du vecteur gaussien  $[\mathbf{p} \mathbf{y}]$  est

$$d(\mathbf{p}, \mathbf{y} | \bar{\mathbf{p}}, \bar{\mathbf{x}}) = K e^{-\frac{1}{2} \bar{\mathbf{z}}^T \mathbf{C}_{[\mathbf{y} \mathbf{p}]}^{-1} \bar{\mathbf{z}}}, \quad \bar{\mathbf{z}} = \begin{pmatrix} \mathbf{y} - F(\bar{\mathbf{p}}, \bar{\mathbf{x}}) \\ \mathbf{p} - \bar{\mathbf{p}} \end{pmatrix} \quad (31)$$

avec  $K$  une constante. Alors, l'EMV  $f_t^n(\mathbf{p}, \mathbf{y})$  de  $[\bar{\mathbf{p}} \bar{\mathbf{x}}]$  est

$$\operatorname{argmin}_{[\bar{\mathbf{p}} \bar{\mathbf{x}}]} \bar{\mathbf{z}}^T \mathbf{C}_{[\mathbf{y} \mathbf{p}]}^{-1} \bar{\mathbf{z}} \text{ avec } \bar{\mathbf{z}} = \begin{pmatrix} \mathbf{y} - F(\bar{\mathbf{p}}, \bar{\mathbf{x}}) \\ \mathbf{p} - \bar{\mathbf{p}} \end{pmatrix}. \quad (32)$$

La fonction  $f_t^n$  définit les estimations  $\mathbf{p}_t^n$  et  $\mathbf{x}_t^n$  de  $\bar{\mathbf{p}}_t$  et  $\bar{\mathbf{x}}_t$  (on note que l'Eq. 18 est un cas particulier de l'Eq. 32).

### 5.5 Propagation de l'erreur (nouvel AFL)

La fonction  $f_t^n(\mathbf{p}_t, \mathbf{y}_t)$  dans l'Eq. 32 peut être réécrite

$$f_t^n(\mathbf{y}_t^n) = \operatorname{argmin}_{\bar{\mathbf{x}}_t^n} \|\mathbf{y}_t^n - F_t^n(\bar{\mathbf{x}}_t^n)\|^2 \quad (33)$$

$$\text{avec } \bar{\mathbf{x}}_t^n = [\bar{\mathbf{p}}_t | \bar{\mathbf{x}}_t], \quad \mathbf{y}_t^n = \mathbf{C}_{[\mathbf{y}_t \mathbf{p}_t]}^{-\frac{1}{2}} \begin{pmatrix} \mathbf{y}_t^T & \mathbf{p}_t^T \end{pmatrix}^T \quad (34)$$

$$\text{et } F_t^n(\bar{\mathbf{x}}_t^n) = \mathbf{C}_{[\mathbf{y}_t \mathbf{p}_t]}^{-\frac{1}{2}} \begin{pmatrix} F_t(\bar{\mathbf{p}}_t, \bar{\mathbf{x}}_t)^T & \bar{\mathbf{p}}_t^T \end{pmatrix}^T. \quad (35)$$

La covariance  $\mathbf{C}_{[\mathbf{y}_t \mathbf{p}_t^n \mathbf{x}_t^n]}$  est nécessaire pour la récurrence. L'Eq. 34 implique  $\mathbf{y}_t^n \sim \mathcal{N}(\bar{\mathbf{y}}_t^n, \mathbf{I})$ ,  $\mathbf{y}_t = \mathbf{K} \mathbf{y}_t^n$  et  $\bar{\mathbf{y}}_t = \mathbf{K} \bar{\mathbf{y}}_t^n$  avec  $\mathbf{K}$  un bloc supérieur de  $\mathbf{C}_{[\mathbf{y}_t \mathbf{p}_t]}^{\frac{1}{2}}$ . On approxime  $f_t^n$  par son développement linéaire de Taylor au point  $\bar{\mathbf{y}}_t^n$  et on obtient

$$\begin{pmatrix} \mathbf{y}_t \\ f_t^n(\mathbf{y}_t^n) \end{pmatrix} = \begin{pmatrix} \bar{\mathbf{y}}_t \\ f_t^n(\bar{\mathbf{y}}_t^n) \end{pmatrix} + \begin{pmatrix} \mathbf{K} \\ \frac{\partial f_t^n}{\partial \bar{\mathbf{y}}_t^n} \end{pmatrix} (\mathbf{y}_t^n - \bar{\mathbf{y}}_t^n). \quad (36)$$

Alors le vecteur  $[\mathbf{y}_t f_t^n(\mathbf{y}_t^n)]$  est Gaussien de covariance

$$\mathbf{C}_{[\mathbf{y}_t \mathbf{p}_t^n \mathbf{x}_t^n]} = \mathbf{C}_{[\mathbf{y}_t f_t^n(\mathbf{y}_t^n)]} = \begin{pmatrix} \mathbf{K} \\ \frac{\partial f_t^n}{\partial \bar{\mathbf{y}}_t^n} \end{pmatrix} \begin{pmatrix} \mathbf{K} \\ \frac{\partial f_t^n}{\partial \bar{\mathbf{y}}_t^n} \end{pmatrix}^T. \quad (37)$$

et  $\frac{\partial f_t^n}{\partial \bar{\mathbf{y}}_t^n}$  est estimé avec les Eqs. 11 et 35.

### 5.6 Algorithme (nouvel AFL)

Les parties 5.4 et 5.5 définissent la relation de récurrence pour le nouvel AFL (Eq. 32). On estime  $[\mathbf{y}_t \mathbf{p}_t \mathbf{x}_t^c]$  et  $\mathbf{C}_{[\mathbf{y}_t \mathbf{p}_t \mathbf{x}_t^c]}$  depuis  $[\mathbf{y}_{t-1} \mathbf{p}_{t-1} \mathbf{x}_{t-1}^c]$  et  $\mathbf{C}_{[\mathbf{y}_{t-1} \mathbf{p}_{t-1} \mathbf{x}_{t-1}^c]}$  comme suit :

1.  $\mathbf{p}_t$  est un sous-vecteur de  $[\mathbf{y}_{t-1} \mathbf{p}_{t-1} \mathbf{x}_{t-1}^c]$
2.  $\mathbf{C}_{[\mathbf{y}_t \cap \mathbf{y}_{t-1} | \mathbf{p}_t]}$  est une sous-matrice de  $\mathbf{C}_{[\mathbf{y}_{t-1} \mathbf{p}_{t-1} \mathbf{x}_{t-1}^c]}$
3. on estime  $\mathbf{C}_{[\mathbf{y}_t \mathbf{p}_t]}$  en utilisant l'Eq. 28
4. on estime  $[\mathbf{p}_t^n \mathbf{x}_t^n] = f_t^n(\mathbf{p}_t, \mathbf{y}_t)$  en utilisant l'AFL (Eq. 32)
5. on estime  $\mathbf{C}_{[\mathbf{y}_t \mathbf{p}_t^n \mathbf{x}_t^n]}$  en utilisant les Eqs. 37, 11 et 35
6. on fait  $\mathbf{p}_t \leftarrow \mathbf{p}_t^n$ ,  $\mathbf{x}_t \leftarrow \mathbf{x}_t^n$  et  $\mathbf{C}_{[\mathbf{y}_t \mathbf{p}_t \mathbf{x}_t]} \leftarrow \mathbf{C}_{[\mathbf{y}_t \mathbf{p}_t^n \mathbf{x}_t^n]}$
7.  $\mathbf{C}_{[\mathbf{y}_t \mathbf{p}_t \mathbf{x}_t^c]}$  est un bloc haut-gauche de  $\mathbf{C}_{[\mathbf{y}_t \mathbf{p}_t \mathbf{x}_t]}$ .

## 6 Expériences

### 6.1 Intégration de la covariance dans la méthode SfM

Notre système SfM temps réel [10] a deux étapes : l’initialisation et la reconstruction incrémentale. La première estime la pose des caméras et des points 3D à partir du début de la séquence en utilisant l’AFG standard. La seconde reconstruit itérativement la séquence (poses et points 3D) en utilisant l’AFL original.

Puis nous avons intégré nos méthodes de calcul de covariance incrémental au système SfM. Ces méthodes ont aussi besoin de la covariance pour les poses de caméra au début de la séquence. Cette covariance est estimée par la méthode standard dérivée de l’AFG [8] : l’inverse de l’approximation du hessien au minimum de la fonction de coût multiplié par le bruit dans l’image  $\sigma$ .

Une jauge simple est choisie pour estimer la covariance des poses du début : on fixe la première pose  $(R_0, t_0)$  avec la rotation  $R_0 = I$  et la position  $t_0 = 0$  et la plus grande coordonnée de la  $t_0^{\text{ème}}$  position  $t_{t_0}$  avec  $t_{t_0}^z = 1$  ( $t_0 = 9$ ). Cette information doit être donnée, car il est connu que la forme des ellipsoïdes d’incertitude provenant des covariances dépend fortement du choix de la jauge [11]. Puis nous retirons les lignes et les colonnes correspondant à ces 7 paramètres dans l’approximation du hessien avant de l’inverser.

### 6.2 Comment vérifier les covariances de l’AFL ?

Nos travaux antérieurs montrent que la géométrie estimée par le SfM basé sur un AFL est proche de celle obtenue par un AFG [10]. On s’attend donc à obtenir des résultats similaires en termes de covariance : la covariance provenant d’un AFL (notre méthode) doit être similaire à celle d’un AFG (méthode standard). On va donc comparer les deux. L’étape finale d’un SfM basé sur l’AFG est un AFG sur toute la séquence. Le vecteur  $\mathbf{x}$  contient tout les paramètres 3D et minimise la fonction de coût  $\mathbf{x} \mapsto \|\mathbf{y} - F(\mathbf{x})\|^2$ , avec  $\mathbf{y}$  le vecteur contenant tout les points 2D suivis au cours de la séquence et  $F$  la concaténation des fonctions de projection correspondantes. Si  $\mathbf{y} \sim \mathcal{N}(\bar{\mathbf{y}}, \sigma^2 \mathbf{I})$ , on obtient une covariance provenant de l’AFG  $\mathbf{C}_{\mathbf{x}} = \sigma^2 (\frac{\partial F}{\partial \mathbf{x}})^T \frac{\partial F}{\partial \mathbf{x}})^{-1}$  [8]. C’est l’inverse de l’approximation du hessien avec les mêmes contraintes de jauge que l’AFG fait sur le début de la séquence (partie 6.1). La valeur de  $\sigma$  est également la même. On réduit le temps de calcul en effectuant l’inversion par bloc seulement pour les paramètres qui nous intéressent [8, 11].

### 6.3 Résultats

La figure 1 montre trois images de la séquence prise en ville. La caméra est calibrée, et est montée derrière le pare-brise du véhicule.

Le véhicule parcourt environ 400m et la séquence contient 2731 images 512. 384 images clefs ont été sélectionnées, 16365 points ont été reconstruits à partir de 74236 points



FIG. 1 – trois images de la séquence.

de Harris [7] mis en correspondances grâce au descripteur SURF [1] (la méthode originale de SURF a été modifiée et ré-implémentée sur GPU en CUDA pour obtenir des performances temps-réel). Les images contiennent en moyenne 193 points suivis sur environ 4.5 images consécutives.

La Figure 2 montre la comparaison quantitative des covariances provenant de l’AFL original avec hypothèse faible (cf 5.3) et celles de l’AFG. La même comparaison est effectuée dans la Figure 3 pour la covariance du nouvel AFL décrite en partie 4 (nouvel AFL avec hypothèse d’indépendance forte).

Dans les deux cas, on étudie l’axe principal des ellipses d’incertitudes de la position des caméras  $\mathbf{c}_t^{t-2}$  avec une probabilité de 90% (la  $t-2^{\text{ème}}$  pose est mise à jour aux temps  $t-2, t-1, t$  à cause de la taille de la fenêtre glissante de l’AFL, et nous choisissons l’incertitude de la dernière mise à jour). L’axe des abscisses est le numéro de l’image clef. Le début de la séquence optimisée par AFG est omis dans ces figures.

En haut des figures 2 et 3, on a le rapport des longueurs des grands axes entre l’AFL et l’AFG. On s’aperçoit que le ratio est acceptable (proche de 1.1) pour l’AFL original. Malheureusement, cette méthode n’est pas temps réel. De plus, le ratio pour le nouvel AFL est petit (proche de 0.6) à cause de l’hypothèse d’indépendance forte entre  $\mathbf{p}_t$  et  $\mathbf{y}_t$ . En bas des figures 2 et 3, on a les angles entre les axes principaux de l’AFL et l’AFG. Les angles sont acceptables (petit) pour les deux méthodes.

Nous avons aussi testé l’AFL original avec hypothèse forte (Section 3). Dans ce cas, le rapport des longueurs des grands axes diverge. Pour cette raison, l’hypothèse forte ne doit pas être utilisée avec l’AFL original.

A ce point, la covariance du nouvel AFL avec hypothèse forte est le seul choix possible pour notre application temps réel (bien que son échelle soit trop petite). Soient  $\bar{e}$  et  $\sigma_e$  la moyenne et l’écart type des ratios de la longueur des grands axes entre l’AFL original (hypothèse faible) et le nouvel AFL (hypothèse forte) pour toutes les images clefs de la séquence. On estime  $\bar{e} = 1.82$  et  $\sigma_e = 0.13$ . Comme  $\sigma_e/\bar{e}$  est petit, nous avons décidé d’améliorer la covariance du nouvel AFL en la multipliant par  $\bar{e}^2$ . Maintenant, les longueurs des axes principaux des ellipsoïdes du nouvel AFL sont presque les mêmes que celle de l’AFL original. La figure 4 montre une vue de dessus de la séquence reconstruite avec les ellipsoïdes de la position de la caméra de notre covariance modifiée.

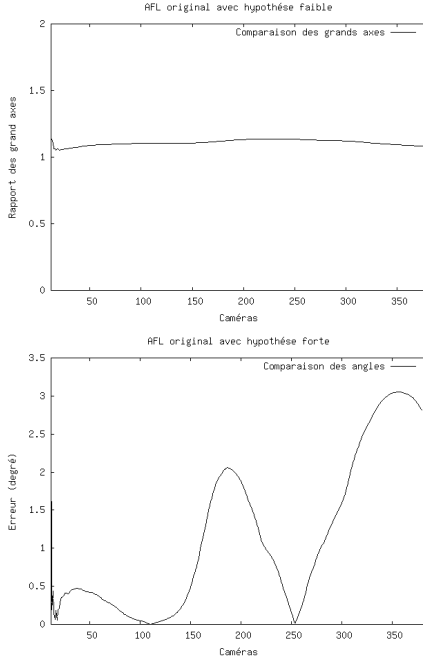


FIG. 2 – Haut : rapport de la longueur des grands axes entre AFL original (hyp. faible) et l’AFG. Bas : angle entre le grand axe de l’AFL original et AFG.

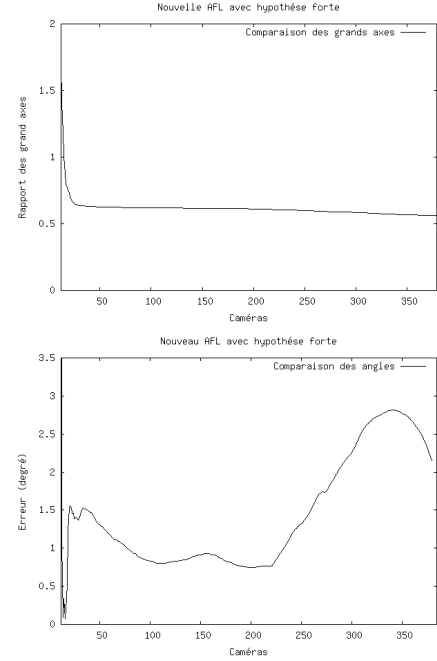


FIG. 3 – Haut : rapport entre la longueur du grand axe entre le nouvel AFL (hyp. forte) et l’AFG. Bas : angle entre le grand axe du nouvel AFL et l’AFG.

Dans la Figure 4, nous voyons que

1. La forme des ellipsoïdes est la même pour le nouvel AFL et l’AFG
2. La longueur du grand axe augmente progressivement avec le temps

Ce sont les résultats attendus.

Seulement 6.4 ms sont nécessaires pour calculer la covariance du nouvel AFL pour chaque image clef. Notre méthode est donc temps réel. Le temps total de calcul de la covariance du nouvel AFL est de 2.4s, qui est très petit devant celui de l’AFG (145s).

Nos expérimentations incluent également des simulations de Monte-Carlo et des estimations de covariance sur d’autres séquences réelles (des résultats similaires ont été obtenus avec des valeurs similaires pour  $\bar{e}$  et  $\sigma_e$ ).

## 7 Conclusion

Cet article propose quatre méthodes d’estimation de la covariance pour des méthodes de Structure-from-Motion (SfM) basées sur l’ajustement de faisceaux local (AFL). Elles proviennent de deux modèles de bruit ("faible" et "forte") et deux ajustements de faisceaux locaux : notre nouvel AFL qui est un Estimateur au Maximum de Vraisemblance pour ces hypothèses et la version originale de l’AFL. Toutes les méthodes propagent l’incertitude tout au long de la séquence, mais seulement deux d’entre elles sont temps réel grâce à l’hypothèse forte.

Nous devons trouver une paire (modèle statistique, estimateur) tel que la covariance estimée soit à la fois plausible

physiquement et temps réel. D’un coté, un modèle statistique réaliste peut demander trop de calcul pour garantir des performances temps réel. D’un autre coté, un modèle statistique non-réaliste peut donner des résultats temps réel mais non conforme à la réalité.

Les expériences effectuées sur les séquences réelles montrent que l’AFL original avec hypothèse faible donne des résultats acceptables. Cette hypothèse est réaliste mais non temps réel. De plus, la méthode originale ne peut être utilisée avec l’hypothèse forte. Le nouvel AFL avec hypothèse forte donne des covariances acceptables en utilisant le coefficient empirique. Pour notre application temps-réel, c’est le seul choix d’estimateur possible.

Des travaux futurs sont la propagation d’erreur des images clefs aux images non clefs, l’intégration des incertitudes sur la calibration dans la propagation d’erreur, des expériences sur les effets du changement de jauge, et la fusion de nos résultats vision avec un GPS ou un odomètre.

## Annexe A

La Proposition 1 (partie 3.2) est un cas particulier de la Proposition 6.1 dans [5]. Cette dernière démontre également que la fonction  $f$  existe localement et est de classe  $\mathcal{C}^1$ .

Dans cette annexe, l’indice  $t$  est omis. De plus,  $F_k$ ,  $\mathbf{x}_i$  et  $\mathbf{p}_j$  sont les  $k^{\text{ème}}$ ,  $i^{\text{ème}}$  et  $j^{\text{ème}}$  coordonnées des vecteurs  $F$ ,  $\mathbf{x}$  et  $\mathbf{p}$ , respectivement.

On élimine  $\mathbf{y}$  de la preuve de l’Eq. 12 car  $\mathbf{y}$  est une constante dans  $F$  pour la dérivation.

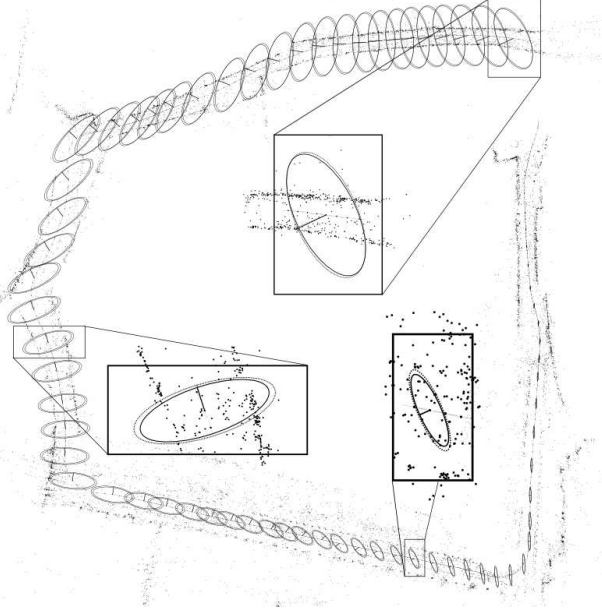


FIG. 4 – Vue de dessus de la reconstruction de la séquence par la méthode SfM basée sur l’AFL. Les positions de caméra et les points 3D sont les points noirs. Pour une caméra sur cinq, les ellipses d’incertitude à 90% des positions sont représentées (trait continu pour l’AFG, trait pointillé pour le nouvel AFL avec hypothèse forte et coefficient)

Les dérivées partielles du second ordre de la fonction

$$g(\mathbf{x}, \mathbf{p}) = \frac{1}{2} \|F(\mathbf{p}, \mathbf{x})\|^2 \quad (38)$$

sont

$$\frac{\partial^2 g}{\partial \mathbf{x}_i \partial \mathbf{p}_j} = \sum_k \left\{ \frac{\partial F_k}{\partial \mathbf{x}_i} \frac{\partial F_k}{\partial \mathbf{p}_j} + \frac{\partial^2 F_k}{\partial \mathbf{x}_i \partial \mathbf{p}_j} F_k \right\}. \quad (39)$$

L’approximation de Gauss-Newton de l’Eq. 39 est

$$\frac{\partial^2 g}{\partial \mathbf{x}_i \partial \mathbf{p}_j} \approx \sum_k \frac{\partial F_k}{\partial \mathbf{x}_i} \frac{\partial F_k}{\partial \mathbf{p}_j} = \left( \frac{\partial F^T}{\partial \mathbf{x}} \frac{\partial F}{\partial \mathbf{p}} \right)_{i,j}. \quad (40)$$

De la même manière, on a

$$\frac{\partial^2 g}{\partial \mathbf{x}_i \partial \mathbf{x}_j} \approx \left( \frac{\partial F^T}{\partial \mathbf{x}} \frac{\partial F}{\partial \mathbf{x}} \right)_{i,j}. \quad (41)$$

Comme  $f(\mathbf{p})$  est le minimum de  $\mathbf{x} \rightarrow g(\mathbf{x}, \mathbf{p})$ , on a

$$\forall i, \frac{\partial g}{\partial \mathbf{x}_i}(f(\mathbf{p}), \mathbf{p}) = 0 \quad (42)$$

Grâce à l’Eq. 42, 40 et 41, on en déduit  $\forall i, j$

$$0 = \frac{\partial}{\partial \mathbf{p}_j} (\mathbf{p} \rightarrow \frac{\partial g}{\partial \mathbf{x}_i}(f(\mathbf{p}), \mathbf{p})) \quad (43)$$

$$= \frac{\partial^2 g}{\partial \mathbf{x}_i \partial \mathbf{p}_j} + \sum_k \left( \frac{\partial^2 g}{\partial \mathbf{x}_i \partial \mathbf{x}_k} \right) \frac{\partial f_k}{\partial \mathbf{p}_j} \quad (44)$$

$$\approx \left( \frac{\partial F^T}{\partial \mathbf{x}} \frac{\partial F}{\partial \mathbf{p}} \right)_{i,j} + \sum_k \left( \frac{\partial F^T}{\partial \mathbf{x}} \frac{\partial F}{\partial \mathbf{x}} \right)_{i,k} \left( \frac{\partial f}{\partial \mathbf{p}} \right)_{k,j} \quad (45)$$

Enfin, l’Eq. 45 est équivalente à

$$\left( \frac{\partial F^T}{\partial \mathbf{x}} \frac{\partial F}{\partial \mathbf{x}} \right) \frac{\partial f}{\partial \mathbf{p}} \approx - \frac{\partial F^T}{\partial \mathbf{x}} \frac{\partial F}{\partial \mathbf{p}}. \quad (46)$$

On obtient l’Eq. 12 car  $\frac{\partial F}{\partial \mathbf{x}}$  est de rang plein. L’Eq. 11 est obtenue comme cas particulier avec  $F(\mathbf{p}, \mathbf{x}) = F(\mathbf{x}) - \mathbf{p}$  en échangeant les notations  $\mathbf{y}$  et  $\mathbf{p}$ .

Le vecteur  $\mathbf{p}_t$  fixe la jauge (pose et échelle) pour l’AFL car  $\mathbf{p}_t$  contient au moins 2 poses [10]. Dans ce contexte,  $\frac{\partial F}{\partial \mathbf{x}}$  est de rang plein pour des configurations générales de points 3D [8].

**Remerciements** Ce travail est financé par le CNRS, le CEA, l’ANR et Num@tec Automotive dans le contexte du projet ODIAAC. Nous remercions E. Mouragnon, S. Lion et S. Naudet pour leurs améliorations du logiciel de reconstruction 3D (SfM), et G. Jacob pour l’implémentation GPU de SURF et Harris en CUDA.

## Références

- [1] H. Bay, T. Tuytelaars, and L. Gool. Surf : Speeded up robust features. In *ECCV’06*.
- [2] C. Bedder and R. Steffen. Incremental estimation without specifying a-priori covariance matrices for the novel parameters. In *VLMP Workshop’08*.
- [3] D. P. Bertsekas. Incremental least squares methods and the extended kalman filter. *SIAM Journal of Optimization*, 6(3), 1996.
- [4] A. J. Davison. Real-time simultaneous localization and mapping with a single camera. In *ICCV’03*.
- [5] O. Faugeras, Q. Long and T. Papadopoulos. *Geometry of Multiple Images*. MIT Press, 2000.
- [6] M. A. Fischler and R. C. Bolles. Random sample consensus : a paradigm for model fitting with application to image analysis and automated cartography. *Communications of the ACM*, 24, 1981.
- [7] C. Harris and M. Stephens. A combined corner and edge detector. In *4th Alvey Vision Conference*, 1988.
- [8] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.
- [9] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *IJCV*, 60, 2004.
- [10] E. Mouragnon, M. Lhuillier, M. Dhome, F. Dekeyser, and P. Sayd. Real time localization and 3d reconstruction. In *CVPR’06*.
- [11] B. Triggs, P. F. McLauchlan, R. Hartley, and A. W. Fitzgibbon. Bundle adjustment – a modern synthesis. In *LNCS*, 2000.