

Un modèle de projection «Rolling Shutter» flexible appliqué au calcul de pose

A flexible rolling shutter projection model applied to pose computation

Ludovic Magerand

Adrien Bartoli

Omar Ait-Aider

LASMEA – Clermont-Ferrand

Résumé

Les caméras «Rolling Shutter», où l’acquisition se fait ligne par ligne, provoquent des distorsions dans l’image lorsqu’un objet rigide se déplace dans la scène, ce qui induit en erreur les modèles classiques de projection. Si sa forme est connue [1] et [2] ont montrés comment retrouver simultanément sa pose et sa vitesse, assumée constante durant l’exposition, à partir d’une seule image.

Nous proposons une méthode flexible pour réaliser le calcul de la pose ligne par ligne dans de telles images et en levant l’hypothèse de vitesse constante. Cela n’est possible qu’avec l’ajout d’un terme de régularisation, pondéré d’un coefficient, au terme d’erreur de reprojection. Afin qu’il ait une réalité géométrique, ce terme est construit à partir de la dérivé du mouvement d’un certain ordre.

La résolution de ce modèle peut être réalisée par la méthode de Levenberg-Marquardt [7] adaptée à la présence de rotations dans les paramètres. L’introduction du coefficient impose l’utilisation d’un critère pour sa sélection comme la norme L-Tangente [3].

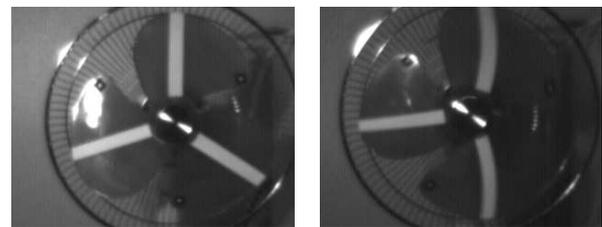
Mots-clefs rolling shutter, distorsion, pose

Abstract

«Rolling Shutter» cameras, where acquisition is done line by line, create distortions in the image when a rigid object moves in the scene, making classical projection model misbehaving. If its shape is known [1] and [2] show how to find, from a single view, both its initial pose and velocity, which is assumed to remain constant during exposition.

We propose a flexible method to compute the pose line by line within such images and removing the constant velocity assumption. This is made possible by adding a weighed regularization term to the reprojection error. To have a geometrical consistency, this term is built from the derivative of the movement at some order.

Resolution of this model is possible by the Levenberg-



(a) Objet au repos.

(b) Objet en mouvement.

FIGURE 2 – Caméras «Rolling Shutter» et distorsions.

Marquardt method [7] adapted to the presence of rotations in the parameters. The weight of the regularization term can be determine by a criterion like the L-Tangent norm [3].

Keywords rolling shutter, distorsion, pose

1 Introduction

Contrairement aux capteurs de technologie CCD, les capteurs à base de technologie CMOS sont peu onéreux et permettent de réaliser des caméras où l’acquisition de l’image se fait ligne par ligne de manière non simultanée, comme illustré figure 1.

Ce mode particulier d’acquisition, appelé «Rolling Shutter», est à l’origine de distorsions dans les images issues de ces caméras lorsqu’un objet se déplace dans la scène filmée. Selon la nature de son mouvement il peut en effet apparaître courbé ou penché sur l’image résultante.

Un exemple d’image présentant une telle distorsion due à un mouvement de rotation est montré figure 2(b), à comparer avec la figure 2(a) qui montre le même objet sans mouvement.

Le mouvement étant à l’origine de ces distorsions, il est donc possible de le caractériser à partir de l’acquisition d’une seule image de ce type. Pour ce faire, il est néces-

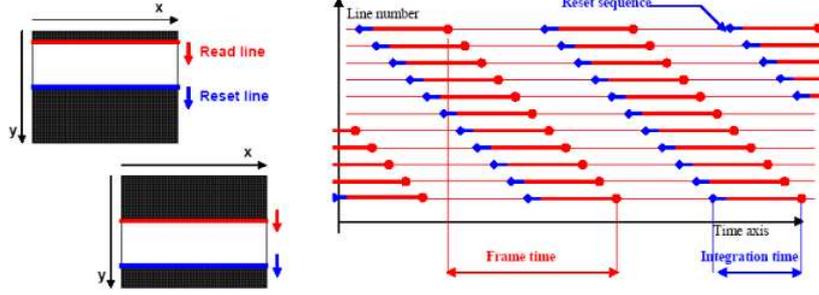


FIGURE 1 – Principe d’acquisition «Rolling Shutter» (extrait de la documentation Silicon Imaging).

saire d’utiliser un *a priori* sur la nature du mouvement afin que le problème soit bien posé.

En assumant que la vitesse de l’objet est constante durant l’acquisition et que sa géométrie est connue [1] et [2] ont proposé des méthodes pour retrouver la pose initiale de l’objet, c’est à dire lors de l’acquisition de la première ligne, ainsi que sa vitesse à partir de points ou de lignes d’intérêts présents dans une seule image et mis en correspondances avec les points 3D de l’objet.

L’approche développée ici consiste à paramétrer la pose ligne par ligne et à utiliser un terme de régularisation pondéré comme contrainte supplémentaire. Afin qu’il ait une interprétation géométrique en rapport avec le mouvement, ce terme est construit à partir de la dérivée du mouvement d’un ordre donné. Celle-ci est obtenue par différence finie à partir de la projection de l’objet par la pose de différentes lignes consécutives autour de la ligne courante. Ce modèle est donc plus flexible et permet de lever l’hypothèse d’uniformité de la vitesse selon l’ordre de dérivation utilisé.

La résolution de ce modèle nécessite de réaliser la minimisation d’un critère non-linéaire qui peut être faite avec une méthode comme Levenberg-Marquardt [7] adaptée à la présence de rotations dans les paramètres. Celle-ci peut se faire de manière efficace car la matrice Jacobienne du vecteur paramètres est bande diagonale dont on sait calculer rapidement la décomposition de Cholesky pour l’inverser. La présence d’un terme de régularisation pondéré nécessite l’évaluation du coefficient de pondération. Cela est fait par l’utilisation du critère de la norme L-Tangente [3] qui fournit une approximation du score de validation croisée [6].

Plan. La section 2 présente les travaux déjà effectués dans la modélisation des caméras perspectives, dont les caméras «Rolling Shutter». La section 3 montre le modèle «Rolling Shutter» flexible, tandis que la section 4 se focalise sur une méthode de résolution de celui-ci. Finalement la section 5 montre quelques résultats obtenus avec cette méthode.

Notations. Les scalaires et indices sont notés en italique (*c*). Les vecteurs sont des matrices colonnes notés en gras (**V**), les autres matrices étant notées en caractères droits (**M**). La transposée d’une matrice **M** est notée \mathbf{M}^T . La ma-

trice identité de taille $n \times n$ est notée \mathbf{I}_n .

2 Travaux existants

Le mode d’acquisition «Rolling Shutter» des capteurs à base de technologie CMOS n’est pas le seul mode existant. Il est en effet possible de les utiliser dans un mode d’acquisition «Global Shutter», qui expose l’ensemble des lignes simultanément. Ce mode d’acquisition produit des images sans distorsions, semblables à celles issues des capteurs de technologie CCD. Dans cette section nous allons présenter quelques uns des résultats importants en modélisation de ces deux modes d’acquisition.

2.1 Les caméras «Global Shutter»

Dans le cas de l’acquisition «Global Shutter», la projection perspective est modélisée sous la forme suivante, comme décrite dans [5] et [9] :

$$s\tilde{\mathbf{m}}_i = \mathbf{K}(\mathbf{R} | \mathbf{t}) \tilde{\mathbf{P}}_i, \quad (1)$$

où $\tilde{\mathbf{P}}_i = (x_i, y_i, z_i, 1)^T$ sont les coordonnées homogènes d’un point 3D i de l’objet dans son repère, et qui se projette en $\tilde{\mathbf{m}}_i = (u_i, v_i, 1)^T$ coordonnées homogènes dans le repère image. s est un facteur d’échelle lié à la profondeur. Les matrices **R** et **t** sont respectivement la rotation et la translation 3D qui représentent la pose du repère objet dans le repère caméra. La matrice **K** est la matrice de projection perspective permettant de réaliser le passage du repère caméra au repère image. En considérant les pixels carrés, elle s’écrit

$$\mathbf{K} = \begin{bmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{bmatrix},$$

où les paramètres α_u et α_v correspondent respectivement aux facteurs d’échelle horizontaux et verticaux. u_0 et v_0 sont les coordonnées de la projection du centre optique dans l’image. Une méthode comme [11] peut être utilisée pour réaliser la calibration de la caméra et connaître ces paramètres.

Si la géométrie de l’objet est connue ainsi que les correspondances entre les points projetés et les points 3D de l’objet, il ne reste comme inconnues que **R**, **t** et s . Le paramètre s peut être alors être éliminé pour permettre de retrouver **R**

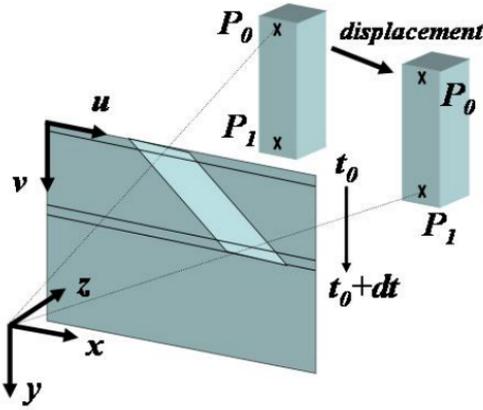


FIGURE 3 – Projection «Rolling Shutter» uniforme.

et \mathbf{t} à partir de l'erreur de reprojection définie pour chaque point par le système

$$\begin{cases} \xi_i^u = u_i - \alpha_u \frac{R^1 \mathbf{P}_i + \mathbf{t}^1}{R^3 \mathbf{P}_i + \mathbf{t}^3} + u_0 \\ \xi_i^v = v_i - \alpha_v \frac{R^2 \mathbf{P}_i + \mathbf{t}^2}{R^3 \mathbf{P}_i + \mathbf{t}^3} + v_0 \end{cases}, \quad (2)$$

où $\mathbf{P}_i = (x_i, y_i, z_i)^\top$, R^j est j -ième ligne de R et \mathbf{t}^j est la j -ième valeur de \mathbf{t} .

En considérant l'ensemble des n points de l'objet, les valeurs recherchées pour les paramètres R et \mathbf{t} sont alors données par

$$\min_{R, \mathbf{t}} \sum_{i=1}^n (\|\xi_i^u\|^2 + \|\xi_i^v\|^2), \quad (3)$$

sous la contrainte que R soit une matrice de rotation.

2.2 Les caméras «Rolling Shutter»

Afin de modéliser les caméras «Rolling Shutter», [1] modifie le modèle de projection perspectif «Global Shutter» en y introduisant une dépendance à la ligne où un point de l'objet est projeté.

Cela est réalisé en considérant que l'objet a une vitesse constante durant l'exposition. Elle est modélisée par une vitesse de translation représentée par le vecteur $\mathbf{V} = (v_1, v_2, v_3)^\top$ et une vitesse de rotation Ω autour d'un axe représenté par le vecteur unitaire $\mathbf{a} = (a_1, a_2, a_3)^\top$.

L'équation (1) de projection du modèle perspectif «Global Shutter» devient alors

$$s\tilde{\mathbf{m}}_i = K (\delta R_i R_0 | \mathbf{t}_0 + \delta \mathbf{t}_i) \tilde{\mathbf{P}}_i,$$

où R_0 et \mathbf{t}_0 représentent la pose initiale de l'objet, c'est à dire à l'instant d'acquisition de la première ligne. δR_i et $\delta \mathbf{t}_i$ représente le mouvement de l'objet entre la première ligne et la ligne v_i . Ce modèle de projection est illustré par la figure 3.

Le temps τ entre l'acquisition de deux lignes consécutives étant connu et la vitesse étant considérée constante, la formule de Rodrigues permet de modéliser la rotation de l'objet entre la première ligne et la ligne v_i par

$$\delta R_i = \mathbf{a} \mathbf{a}^\top (1 - \cos(\tau v_i \Omega)) + I_3 \cos(\tau v_i \Omega) + \hat{\mathbf{a}} \sin(\tau v_i \Omega).$$

La translation entre ces deux mêmes lignes est quand à elle donnée par

$$\delta \mathbf{t}_i = \tau v_i \mathbf{V}.$$

L'erreur de reprojection de ce modèle s'obtient depuis l'équation (2) en remplaçant R par $R_i = \delta R_i R_0$ et \mathbf{t} par $\mathbf{t}_i = \mathbf{t}_0 + \delta \mathbf{t}_i$:

$$\begin{cases} \xi_i^u = u_i - \alpha_u \frac{R_i^1 \mathbf{P}_i + \mathbf{t}_i^1}{R_i^3 \mathbf{P}_i + \mathbf{t}_i^3} + u_0 \\ \xi_i^v = v_i - \alpha_v \frac{R_i^2 \mathbf{P}_i + \mathbf{t}_i^2}{R_i^3 \mathbf{P}_i + \mathbf{t}_i^3} + v_0 \end{cases}.$$

Les paramètres de cette erreur sont alors Ω , \mathbf{a} , \mathbf{V} , R_0 et \mathbf{t}_0 . La minimisation présentée équation (3) devient donc

$$\min_{\Omega, \mathbf{a}, \mathbf{V}, R_0, \mathbf{t}_0} \sum_{i=1}^n (\|\xi_i^u\|^2 + \|\xi_i^v\|^2),$$

sous la contrainte que \mathbf{a} soit unitaire et R_0 une matrice de rotation.

3 Modèle «Rolling Shutter» flexible

Le modèle présenté à la section 2.2 permet de retrouver simultanément la pose et la vitesse de l'objet ainsi que de corriger les distorsions induites par le mode d'acquisition «Rolling Shutter». Il serait néanmoins intéressant de lever la contrainte d'uniformité de la vitesse lors de l'acquisition. Pour cela une modélisation flexible de la projection par une caméra «Rolling Shutter» va être développée dans cette section.

3.1 Modèle de projection

Le modèle de projection perspectif présenté équation (1) peut être modifié pour permettre la prise en compte des effets de distorsion de la manière suivante :

$$s\tilde{\mathbf{m}}_i = K (R_{v_i} | \mathbf{t}_{v_i}) \tilde{\mathbf{P}}_i,$$

où R_{v_i} et \mathbf{t}_{v_i} représentent la pose de l'objet à l'instant de l'acquisition de la ligne v_i .

Ce principe de projection est illustré à la figure 4. Les lignes en pointillés longs symbolisent les lignes de l'image, et les lignes en pointillés courts la projection. Trois points connus de l'objet (\mathbf{P}_1 à \mathbf{P}_3) sont projetés sur l'image, chacun avec la pose de la ligne correspondante. Il faut remarquer qu'il peut exister des lignes où aucun point connu de l'objet n'est projeté, comme c'est le cas de la troisième ligne (\mathbf{P}_3 se projette sur la quatrième ligne).

Pour les lignes comportant un des points projetés, les composantes de l'erreur de reprojection sont les mêmes que celles de l'équation (2) en substituant R_{v_i} à R et \mathbf{t}_{v_i} à \mathbf{t} :

$$\begin{cases} \xi_{v_i}^u = u_i - \alpha_u \frac{R_{v_i}^1 \mathbf{P}_i + \mathbf{t}_{v_i}^1}{R_{v_i}^3 \mathbf{P}_i + \mathbf{t}_{v_i}^3} + u_0 \\ \xi_{v_i}^v = v_i - \alpha_v \frac{R_{v_i}^2 \mathbf{P}_i + \mathbf{t}_{v_i}^2}{R_{v_i}^3 \mathbf{P}_i + \mathbf{t}_{v_i}^3} + v_0 \end{cases}. \quad (4)$$

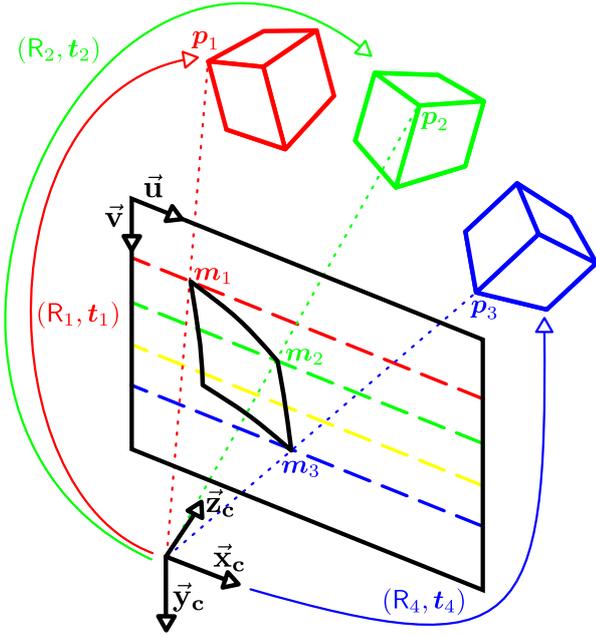


FIGURE 4 – Projection «Rolling Shutter» flexible.

Ce modèle de projection permet de modéliser tout type de mouvement mais sa résolution uniquement avec l'erreur de reprojection est un problème mal posé. En effet il n'existe aucun lien entre les couples (R_{v_i}, t_{v_i}) représentant la pose des différentes lignes. De plus l'erreur de reprojection n'est pas définie sur les lignes ne comportant pas de point projeté. Ce modèle a donc plus d'inconnues que de contraintes, ce qui aboutit à un ensemble de plusieurs solutions possibles.

Il est donc nécessaire d'ajouter une composante supplémentaire sous la forme d'un terme de régularisation ξ_k^r où $k \in [1, l]$, l étant le nombre total de lignes consécutives considérées de l'image et contenant à minima toute les lignes où un point de l'objet est projeté. Les paramètres recherchés sont alors $\{(R_k, t_k)\}_{k=1}^l$ et sont donnés par la minimisation :

$$\min_{\{(R_k, t_k)\}} \sum_{i=1}^n (\|\xi_{v_i}^u\|^2 + \|\xi_{v_i}^v\|^2) + \frac{1-\beta}{\beta} \sum_{k=1}^l (\|\xi_k^r\|^2), \quad (5)$$

sous la contrainte que les matrices R_k soient des matrices de rotation. $\beta \in [0, 1]$ est un coefficient de pondération nécessaire pour prendre en compte le fait que le terme de régularisation comporte davantage de contraintes que l'erreur de reprojection. Sa détermination fera l'objet de la section 4.3.

3.2 Terme de régularisation

Le terme de régularisation est construit à partir de la dérivée du mouvement d'ordre d projetée sur le plan image. Celle-ci est obtenue à l'aide d'un schéma de différence finie centrale appliqué à la projection 2D d'un point 3D P_i de l'objet par la pose des d lignes autour de la ligne k consi-

dérée par

$$\delta^d \psi(P_i, k) = \sum_{p=0}^d (-1)^p \binom{d}{p} \psi(P_i, k + \left\lfloor \frac{d}{2} \right\rfloor - p),$$

où $\psi(P_i, k)$ représente la projection 2D du points 3D P_i de l'objet par la pose de la ligne k .

Soit $Q_i = (q_i^1, q_i^2, q_i^3)^\top$ le point 3D résultat de la transformation du point P_i par la pose de la ligne k définis par

$$Q_i = K(R_k | t_k) P_i.$$

La projection 2D $\psi(P_i, R_k, t_k)$ de ce point est alors

$$\psi(P_i, k) = \begin{pmatrix} q_i^1 \\ q_i^2 \\ q_i^3 \end{pmatrix}^\top.$$

Afin que le terme de régularisation rende bien compte du mouvement de l'ensemble de l'objet, il est constitué de la somme des dérivées obtenues pour l'ensemble des n points P_i de l'objet :

$$\xi_k^r = \sum_{i=1}^n \delta^d \psi(P_i, k).$$

Une illustration de ce terme au premier ordre est donnée par la figure 5 où il est représenté par les lignes en pointillés alternés courts et longs. Pour plus de clarté, il est supposé sur cette figure que l'on ne connaît que deux points de l'objet et qu'ils se projettent sur deux lignes consécutives. En observant la figure 5, il est possible de déterminer la nature de la contrainte imposée au mouvement : sa vitesse doit être nulle pour minimiser ce critère. En transposant ce raisonnement au second ordre, cette contrainte est que l'accélération doit être nulle, et donc la vitesse constante. Pour le terme d'ordre trois, la contrainte imposée est alors que le mouvement se fait à accélération constante, etc.

Le terme de régularisation a donc une interprétation géométrique et physique forte et peut être adapté à la nature du mouvement de l'objet dans la scène filmée en choisissant l'ordre de dérivation d de manière adéquate. Cela permet une certaine flexibilité de la modélisation de la projection présentée section 3.1.

Le cas du terme d'ordre premier est un peu particulier et peut s'interpréter de deux façons : soit la scène est effectivement filmée par une caméra «Rolling Shutter» auquel cas elle est rigide, soit elle est filmée par une caméra «Global Shutter» et donc le mouvement n'est pas perçu.

4 Résolution du modèle

4.1 Optimisation des paramètres

Le modèle de projection «Rolling Shutter» flexible présenté à la section 3 peut être résolu vis-à-vis de ces paramètres par l'équation (5). Elle fait intervenir la minimisation d'un critère non-linéaire qui peut être réalisée par la méthode de Levenberg-Marquardt [7]. Celle-ci procède

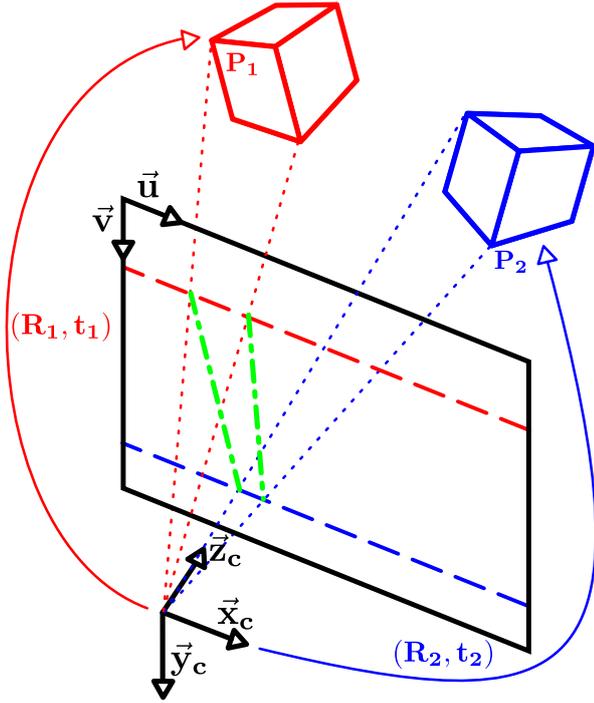


FIGURE 5 – Terme de régularisation au premier ordre.

par itérations successives en mettant à jour le vecteur des paramètres \mathbf{x} :

$$\mathbf{x} \leftarrow \mathbf{x} + \mathbf{d}(\mathbf{x}, \lambda), \quad (6)$$

où λ est un coefficient mis à jour à chaque itération et qui permet de faire converger la méthode soit vers une méthode de descente du gradient dont la convergence est assurée mais lente, soit vers une méthode Gauss-Newton de convergence rapide mais non assurée.

La fonction $\mathbf{d}(\mathbf{x}, \lambda)$ fait intervenir la matrice Jacobienne $J(\mathbf{x})$ comme présenté dans [10] :

$$\mathbf{d}(\mathbf{x}, \lambda) = (J(\mathbf{x})^\top J(\mathbf{x}) + \lambda I_p)^{-1} J(\mathbf{x})^\top \mathbf{r}(\mathbf{x}),$$

où p est le nombre de paramètres du modèle et $\mathbf{r}(\mathbf{x})$ le vecteur des résidus en \mathbf{x} .

En raison de la présence de rotations dans les paramètres de la minimisation, il faut remplacer la mise à jour additive de l'équation (6) par une mise à jour en deux temps : additive pour les paramètres de translation et multiplicative pour les rotations.

Il ne reste qu'à initialiser l'optimisation et pour cela il est possible de prendre le résultat du modèle «Global Shutter» qui donne une bonne approximation de la pose globale, laquelle sera affinée par notre modèle pour chaque ligne en prenant en compte les distorsions dues au mouvement.

4.2 Matrice Jacobienne des paramètres

Le calcul de la matrice Jacobienne des paramètres $J(\mathbf{x})$ et l'inversion de $J(\mathbf{x})^\top J(\mathbf{x}) + \lambda I_p$ peuvent devenir très coûteux en temps de calcul. Il convient donc d'étudier la structure de $J(\mathbf{x})$ pour réaliser ces calculs de manière efficace.

Le vecteur des paramètres, de taille $6l$, peut être constitué selon

$$\mathbf{x} = (\mathbf{t}_1^\top \dots \mathbf{t}_l^\top \mid \mathbf{E}_1^\top \dots \mathbf{E}_l^\top)^\top,$$

où \mathbf{E}_k est le vecteur contenant les angles d'Euler correspondant à \mathbf{R}_k .

Le vecteur des résidus, de taille $2n(l+1)$, peut être construit selon

$$\mathbf{r}(\mathbf{x}) = (\xi_1^u, \xi_1^v, \dots, \xi_n^u, \xi_n^v \mid \xi_1^r, \dots, \xi_l^r)^\top.$$

La matrice Jacobienne est alors de la forme

$$J(\mathbf{x}) = \begin{bmatrix} \mathbf{1} & \mathbf{1} & 0 & \dots & 0 \\ \mathbf{1} & \ddots & \ddots & 0 & \vdots \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & 0 & \ddots & \ddots & \mathbf{1} \\ 0 & \dots & 0 & \mathbf{1} & \mathbf{1} \end{bmatrix},$$

où les blocs $\mathbf{1}$ représentent des blocs de la matrice non vides.

C'est donc une matrice creuse de taille $2n(l+1) \times 6l$ bande diagonale. Il est possible de calculer une décomposition de Cholesky efficace de telles matrices, ce qui permet de calculer $\mathbf{d}(\mathbf{x}, \lambda)$ rapidement.

4.3 Coefficient de pondération

L'équation (5) introduit un coefficient de pondération entre l'erreur de reprojection et le terme de régularisation. La détermination de la valeur de ce coefficient peut se faire à l'aide de plusieurs critères comme le score de validation croisée [6] ou une de ces variations. Ces critères sont difficiles à minimiser et nécessitent un temps de calcul conséquent. Une alternative possible est d'approximer le score de la validation croisée, comme dans [4] par exemple.

D'autres critères basés sur l'utilisation de L-Curve, comme dans [8], peuvent être aussi être utilisés. L'idée de ces scores est de mettre en compétition l'erreur de reprojection totale

$$\xi_d = \beta \sum_{i=1}^n (\|\xi_{v_i}^u\|^2 + \|\xi_{v_i}^v\|^2),$$

avec l'erreur de régularisation totale

$$\xi_r = (1 - \beta) \sum_{k=1}^l (\|\xi_k^r\|^2).$$

La valeur de β choisie est alors celle constituant le meilleur rapport entre ces deux erreurs.

Généralement plusieurs valeurs de β correspondent à ce critère, et le choix de la valeur correcte n'est pas clair. Le critère retenu ici est donc la norme L-Tangente présentée dans [3] et qui fournit une bonne approximation du score de la validation croisée.

La norme L-Tangente permet de choisir la valeur correcte en remarquant qu'autour de celle-ci, un petit changement

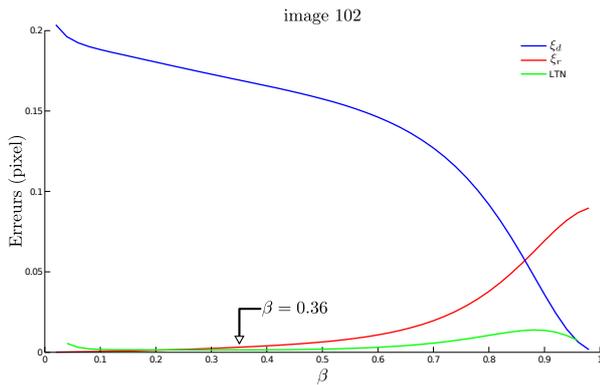


FIGURE 6 – Courbes des erreurs et de la norme L-Tangente (LTN). Le coefficient de pondération est choisit comme étant le minimum le plus favorable (ici $\beta = 0.36$).

dans cette valeur n'implique pas une grosse évolution des deux erreurs.

Cela revient à trouver un minimum de la fonction

$$LTN(\lambda) = \|(\xi'_d, \xi'_r)\|_2^2,$$

où ξ'_d et ξ'_r sont les dérivés de ξ_d et ξ_r par rapport à λ . Ces dérivées peuvent être approximées par un schéma de différence finie.

Cette fonction présente généralement plusieurs minima, dont deux bien souvent situés près des valeurs $\beta = 0$ et $\beta = 1$ qui peuvent être éliminés si un autre minimum est présent.

Afin de s'assurer de ne pas tomber dans un de ces deux minima et pour atteindre le minimum réellement recherché, il suffit d'initialiser le paramètre de la norme L-Tangente à la valeur $\beta = 0.5$.

Un exemple des courbes d'erreurs et de calcul de la norme L-Tangente est donné figure 6, le minimum recherché est dans ce cas situé en $\beta = 0.36$.

5 Résultats

La méthode décrite dans la section 4 a été implémentée en utilisant les ordres de dérivation deux et trois. Elle est comparée aux modèles «Global Shutter» et «Rolling Shutter» uniforme.

Des données simulées ont été générées à partir de la projection de trois faces d'un cube virtuel par une caméra «Rolling Shutter». Quatre simulations d'images ont été réalisées à paramètres cinématiques quadratiques : sans mouvement, avec une translation pure, avec une rotation pure, et enfin avec un mouvement composé. Quatre niveaux de bruit additif gaussien ont été appliqués aux coordonnées : inférieur au pixel, de l'ordre du pixel, de quelques pixels et d'une dizaine de pixels.

Les données réelles sont issues d'une séquence sur laquelle se trouve un objet en rotation plane, aux vibrations près. La séquence comporte 46 images, correspondant aux deux premiers tours de l'objet. Lors du premier tour, la vitesse

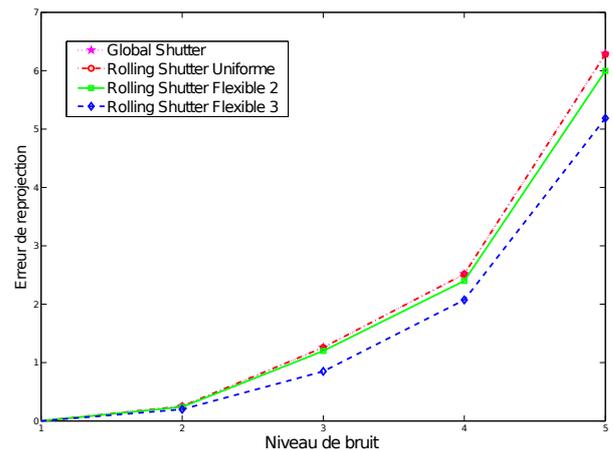


FIGURE 7 – Erreur de reprojection en fonction du bruit appliqué lorsqu'aucun mouvement n'est appliqué à l'objet. L'ordre de grandeur de l'erreur pour les 4 modèles est à peu près équivalent. Les modèles «Global Shutter» et «Rolling Shutter» uniforme sont très proches car bien plus contraints que le modèle «Rolling Shutter» flexible.

reste faible et les distorsions dans les images sont peu importantes. Au deuxième tour la vitesse est plus conséquente et les distorsions sont très marquées.

5.1 Données simulées

Deux critères de comparaison sont retenus. Tout d'abord l'erreur de reprojection telle que obtenue à la fin des optimisations. Ensuite l'erreur entre la reprojection sur une image considérée comme référence (celle sans bruit) et la reprojection avec les paramètres optimisés.

La figure 7 fait apparaître le comportement de l'erreur de reprojection en fonction du niveau de bruit appliqué. Dans le cas d'un objet statique, l'erreur de reprojection est globalement comparable pour les quatre méthodes, avec des valeurs légèrement plus faible pour les modèles «Rolling Shutter» flexible. Cela s'explique par le fait que dans ce modèle l'erreur sur le terme de régularisation est plus facile à réduire que celle sur la reprojection. En effet, l'erreur sur la régularisation peut diminuer en modifiant les poses des lignes voisines sans impacter l'erreur de reprojection si elles ne contiennent aucun point. C'est d'autant plus vrai avec le troisième ordre de dérivation qui permet intrinsèquement plus de liberté.

Sur la figure 8, le mouvement est une composition de rotation et de translation. Il est évident que le modèle «Global Shutter» est inadapté à ces données, son erreur se situant à plus du double de celle des autres modèles. Le modèle «Rolling Shutter» uniforme se comporte mieux mais n'est pas non plus idéal pour des données à cinématique non constante durant l'acquisition d'une image. Les modèles obtenant la meilleure erreur de reprojection, à des niveaux tout à fait comparables, sont les deux ordres de dérivation du modèle «Rolling Shutter» flexible.

Bien que le modèle «Rolling Shutter» flexible s'adapte

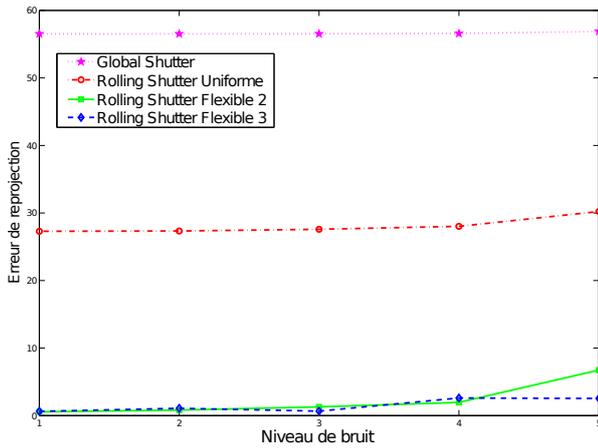


FIGURE 8 – Erreur de reprojection en fonction du bruit appliqué avec un mouvement composé d’une rotation et d’une translation. Le modèle «Global Shutter» est vraiment très loin des autres modèles. Les deux ordres de dérivation du modèle «Rolling Shutter» flexible donne des résultats tout à fait comparables.

mieux aux données, il pourrait néanmoins diverger de la vérité. L’utilisation d’une référence et l’étude de l’erreur par rapport à cette référence permet de réaliser une étude plus complète. Le cas du mouvement de rotation pure présentée figure 9 permet de remarquer à quel point les deux ordres de dérivation du modèle «Rolling Shutter» flexible sont proches. En effet, c’est une des images qui expose la plus grande distinction entre les deux, et ils restent pourtant du même ordre de grandeur.

En figure 10, le mouvement est composé et permet de remarquer un point important : le modèle «Rolling Shutter» flexible se comporte très mal face au bruit au delà d’une certaine limite. À partir du troisième niveau de bruit, l’erreur commise par rapport à la référence redevient plus importante que celle du modèle «Rolling Shutter» uniforme bien que restant inférieure à celle du modèle «Global Shutter». C’est un phénomène qui a été remarqué de façon plus ou moins prononcé sur d’autres tests. Il s’agit là encore probablement d’un effet de bord de la compétition entre le terme d’erreur de reprojection et le terme de régularisation.

5.2 Données réelles

La figure 11 présente l’évolution de la pose de l’objet filmé tout au long de la séquence. Il faut remarquer que le modèle «Global Shutter» est encore une fois complètement inadapté à ce type de données, il diverge en effet très vite du plan de rotation. Les modèles «Rolling Shutter» trouve un plan de rotation qui est bien plus réaliste. Ce qui différencie principalement le modèle «Rolling Shutter» uniforme du modèle «Rolling Shutter» flexible, c’est plus la dispersion autour de se plan que le plan lui-même. En effet, la table 1 fournit l’écart type de la translation selon l’axe de profondeur et confirme bien la stabilité du plan obtenu.

Une brève comparaison a été réalisée sur cette séquence sur

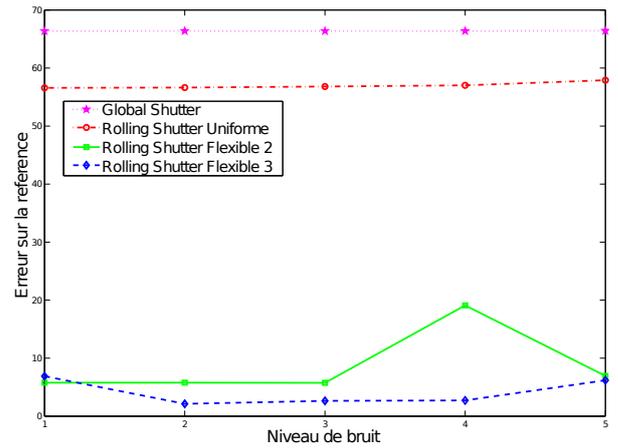


FIGURE 9 – Erreur par rapport à la référence en fonction du bruit appliqué avec un mouvement de rotation. Même lorsque les deux ordres de dérivation du modèle «Rolling Shutter» flexible se distingue vraiment l’un de l’autre, ils restent relativement semblables et plus proche de la référence que les autres modèles.

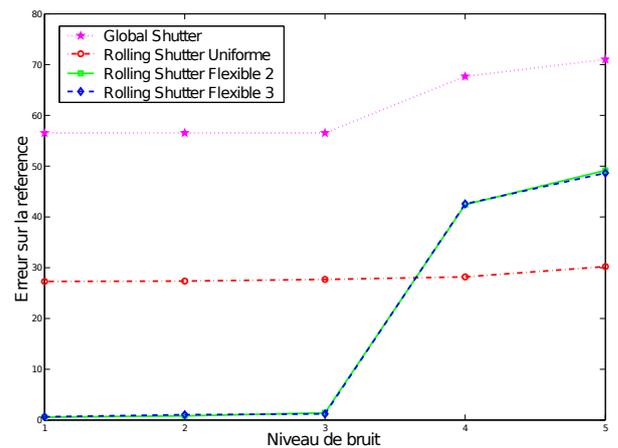


FIGURE 10 – Erreur par rapport à la référence en fonction du bruit appliqué pour un mouvement composé. Le comportement est globalement le même que pour l’erreur de reprojection. Néanmoins les deux modèles «Rolling Shutter» flexible se comportent mal face au bruit lorsque celui-ci devient trop important (de l’ordre de quelques pixels).

	G.S.	R.S.U.	R.S.F.
Moyenne	0,5385	0,5353	0,5338
Écart-type	0,0139	0,0086	0,0035

TABLE 1 – Moyenne et écart-type des trois modèles sur la profondeur de la pose reconstruite.

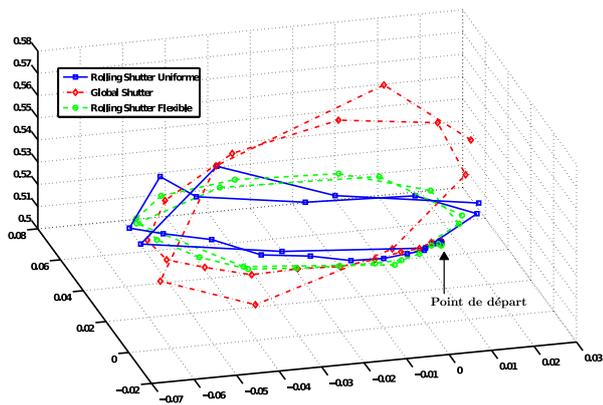


FIGURE 11 – Tracé 3D de l'évolution de la pose d'un objet en rotation. Les deux modèles «Rolling Shutter» sont bien adaptés et trouvent le bon plan moyen de rotation. Le modèle «Rolling Shutter Uniforme» oscille davantage autour de ce plan que le modèle «Rolling Shutter» flexible.

le temps nécessaire pour réaliser l'optimisation du modèle «Rolling Shutter» flexible. À l'ordre trois, il faut jusqu'à sept fois plus d'itération pour atteindre le résultat qu'avec le deuxième ordre. Cela a pour effet de rendre l'ordre trois beaucoup plus long à optimiser.

6 Conclusion

Le modèle développé dans cet article permet de calculer avec précision la pose d'un objet se déplaçant devant une caméra en mode d'acquisition «Rolling Shutter». Il prend en compte les distorsions induites par ce mode et qui permettent par ailleurs de déterminer la pose à chaque ligne de l'image. Par la flexibilité qu'il offre, il permet de faire différentes hypothèses quant au mouvement de l'objet en imposant qu'un certain ordre de dérivation de celui-ci soit nul.

Une comparaison des différents modèles de projection possibles sur des données simulées avec un mouvement non constant ont permis de montrer que le modèle «Rolling Shutter» flexible est plus adapté pour ce type de données. Néanmoins, il faut faire très attention au niveau de bruit sur les coordonnées des points suivis dans l'image car cette méthode est très sensible à un bruit élevé en raison de la compétition entre le terme d'erreur de reprojection et le terme de régularisation.

Sur des images réelles d'un objet en mouvement sujet à une certaine accélération devant une caméra en mode d'acquisition «Rolling Shutter», il a été montré que les modèles «Rolling Shutter» permettent bien de corriger les distorsions. L'erreur de reprojection devient ainsi subpixelique, contrairement au modèle de projection «Global Shutter» qui est induit fortement en erreur et donc inadapté à ce cadre.

La modélisation «Rolling Shutter» flexible apporte aussi un avantage par rapport au modèle uniforme en levant l'hypothèse d'uniformité de la vitesse, ce qui a pour consé-

quence d'améliorer la stabilité des poses reconstruites lorsque celle-ci n'est plus vérifiée.

Il serait d'ailleurs possible de déterminer automatiquement l'ordre de dérivation à utiliser en regardant le comportement des paramètres du modèle, et notamment du coefficient de pondération entre l'erreur de reprojection et le terme de régularisation.

Références

- [1] O. Ait-Aider, N. Andreff, J.-M. Lavest, and P. Martinet. Simultaneous object pose and velocity computation using a single view from a rolling shutter camera. In *European Conference on Computer Vision*, 2006.
- [2] O. Ait-Aider, A. Bartoli, and N. Andreff. Kinematics from lines in a single rolling shutter image. In *Computer Vision and Pattern Recognition*, 2007.
- [3] F. Brunet, A. Bartoli, R. Malgouyres, and N. Navab. L-tangent norm : A low computational cost criterion for choosing regularization weights and its use for range surface reconstruction. In *3D Data Processing Visualization and Transmission*, 2008.
- [4] M. Farenzena, A. Bartoli, and Y. Mezouar. Efficient camera smoothing in sequential structure-from-motion using approximate cross-validation. In *European Conference on Computer Vision*, 2008.
- [5] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2004.
- [6] R. Kohavi. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *International Joint Conferences on Artificial Intelligence*, 1985.
- [7] D. W. Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *Society for Industrial and Applied Mathematics*, 11, 1963.
- [8] H. P. Analysis of discrete ill-posed problems by means of the l-curve. *Society for Industrial and Applied Mathematics*, 34, 1992.
- [9] N. Paragios, Y. Chen, and O. Faugeras. *Handbook of Mathematical Models in Computer Vision*. Springer, 2006.
- [10] B. Triggs, A. Zisserman, and R. Szeliski. *Vision Algorithms : Theory and Practice*. Springer, 1999.
- [11] R. Tsai. An efficient and accurate camera calibration technique for 3d machine vision. *Computer Vision and Pattern Recognition*, 86, 1986.