Estimation de mouvement d'un système stéréoscopique hybride à partir des droites

Sang Ly[1] Cédric Demonceaux[1] Pascal Vasseur[1,2]

[1]Laboratoire MIS EA 4290.Université de Picardie Jules Verne
7 rue du Moulin Neuf. 80000 Amiens.
[2]Laboratoire Heudiasyc. Université de Technologie de Compiègne
Centre de Recherches de Royallieu - BP 20529 - 60205 Compiègne.
{sang.ly, cedric.demonceaux, pascal.vasseur}@u-picardie.fr

Résumé

Nous présentons une méthode d'estimation du déplacement d'un banc stéréoscopique hybride basée sur les droites de la scène. L'approche proposée peut s'appliquer à tous types de systèmes hybrides construits à partir de caméras à point de vue unique telles que perspectives, catadioptriques centrales ou fisheye. Pour ce faire, nous projetons les images sur le modèle unifié de la sphère unité. Sur ces images sphériques, nous pouvons estimer la rotation des caméras à partir des points de fuites de la scène. Puis, nous calculons les translations à partir des projections des droites 3D sur les images sphériques. L'algorithme a été validé aussi bien sur des données de synthèse que sur des images réelles prises à partir d'un banc stéréoscopique muni d'une caméra perspective et d'une caméra fisheye et d'un banc composé de deux caméras para-catadiotpriques.

Mots Clef

Estimation du mouvement, vision stéréoscopique, navigation de robots

Abstract

We present a motion estimation approach for hybrid stereo rigs using line images. The proposed method can be applied to a hybrid system built up from any single view point (SVP) cameras such as perspective, central catadioptric and fisheye cameras. Such configuration combines advantageous characteristics of different types of cameras. Images captured by SVP imaging devices may be mapped to spherical images using the unified projection model. It is possible to recover the camera orientations using vanishing points of parallel line sets. We then estimate the translations from known rotations and line images on the spheres. The algorithm has been validated on simulated data and real images taken from a perspective-fisheye stereo rig and two para-catdioptric stereo rig.This vision-based estimation method can be employed in navigation of autonomous robots besides conventional devices such as Global Positioning System (GPS) and Inertial Navigation System (INS). It helps vision-based localization of robots equiped with stereo vision systems.

Keywords

Motion estimation, stereo vision, robot navigation.

1 Introduction

La localisation est un problème important pour le contrôle d'un robot en déplacement. Cette localisation peut être obtenue à l'aide de différents capteurs. Alors que les pertes de signaux sont fréquents pour les GPS et que les centrales inertielles peuvent accumuler les erreurs relatives au cours du temps, les approches basées vision semblent être un choix judicieux. Remarquons que les caméras à large champ de vue apportent un avantage certain pour la localisation d'un robot mais au détriment d'une faible résolution et de distorsions des images obtenues, à l'inverse les caméras perspectives bénéficient d'une grande résolution mais d'un champ de vue limité. Ainsi, l'utilisation d'un type de caméra dépend fortement de l'application souhaitée. C'est pourquoi nous développons ici une méthode valide pour tous types de caméras à point de vue unique. Il existe de nombreuses méthodes d'estimation du déplacement d'une caméra. Ces méthodes peuvent être classées en trois catégories [4] : 1. estimation de l'egomotion à partir du flot optique, 2. décomposition de la matrice fondamentale, essentielle (F/E) ou homographie (H) 3. estimation en deux phases par un découplage de la rotation et de la translation. Ce papier présente une nouvelle méthode en deux phases pour l'estimation du déplacement d'un banc stéréoscopique hybride.

Les systèmes de vision stéréoscopiques perspectifs sont très largement utilisés dans les problèmes de *structure from motion* (SFM) [16, 22, 24]. Dans [15], Mallet et al. ont estimé les six paramètres de déplacement d'un robot mobile à partir d'une mise en correspondance et de suivi de points le long de la séquence. Olson et al. ont proposé dans [19] une approche robuste du déplacement 3D par maximum de vraisemblance. Dans [13], Lemaire et al. ont présenté deux approches pour le problème du SLAM : une basée

stéréovision et une autre mono-caméra. Ils ont proposé une mise en correspondance robuste des points d'intérêts en tenant compte à la fois des informations photométriques et géométriques des points de Harris. Comport et al. [5] ont développé une technique d'odométrie visuelle 3D afin de suivre la trajectoire d'un véhicule équipé d'un système de stéréovision.

Les systèmes de vision stéréoscopiques omnidirectionnels ont eux aussi été largement développés ces dernières années. Gluckman et al. ont introduit dans [8] un système panoramique stéréoscopique composé de deux capteurs paracatadioptriques (miroir paraboloïde et caméra orthographique) dont les axes optiques sont alignés verticalement. Plus tard, Ollis et al. [18] ont simulé différents systèmes stéréoscopiques catadioptriques avec des miroirs hyperboloïdes et une ou deux caméras. Dans [14], Li a développé un système stéréo temps réel avec des caméras fisheye où une rectification des images en images de latitude et longitude permet d'accélérer la correspondance entre les points.

Les systèmes stéréoscopiques hybrides sont construit à partir de caméras de différentes catégories. Ces configurations nécessitent l'utilisation d'un modèle générique pour étudier les différents types de caméras. Dans [20], Sturm a utilisé une modélisation générique basée sur les lignes de vues. Ce modèle permet d'obtenir une contrainte multifocale pour des caméras de différents classes.

Dans cet article, nous présentons une méthode d'estimation du déplacement d'un banc stéréo composé de deux caméras à point de vue unique, pour laquelle :

1. Nous utilisons le modèle de projection unifié proposé par Mei [17], qui est une version modifiée des modèles développés par Geyer [7] et Barreto [2]. Cette modélisation permet de prendre en compte tout type de caméra à projection centrale et inclus les lentilles fisheye [23]. Ainsi, notre méthode peut s'appliquer aussi bien aux caméras perspective, catadioptrique centrale ou encore fisheye.

2. Seules les lignes sont utilisées comme primitives. Ces primitives sont plus stables que les points et sont moins susceptibles d'être produites par hasard (bruit par exemple), surtout dans des environnements urbains [6]. Comparées aux points d'intérêts, les lignes sont moins nombreuses mais portent plus d'informations, elles ont des informations géométriques et des caractéristiques topologiques très utiles pour la mise en correspondance [3, 9]. De plus, les auteurs dans [1, 4] calculent les rotations à partir des lignes puis les translations par des points d'intérêts. Ici, nous proposons une méthode nécessitant exclusivement l'extraction des droites afin de calculer à la fois les rotations mais aussi les translations, ce qui optimise le temps de calcul des techniques à deux étapes et trouve ainsi son utilité en robotique mobile.

Dans la section suivante, nous développons la géométrie multi-vues pour les caméras à point de vue unique. Puis, nous présentons notre méthode d'estimation du déplacement à partir des lignes. Enfin, nous montrons des résultats expérimentaux sur des données de synthèse et réelles avant de conclure.

2 Géométrie multi-vues pour caméra à point de vue unique

Les caméras centrales et fisheye peuvent être modélisées par une sphère unitaire, et sont ainsi toutes considérées comme des caméras sphériques. Notons que pour connaître le déplacement d'une caméra à partir de lignes, au moins trois vues sont nécessaires [10], c'est pourquoi nous développons ici la géométrie trois et quatre vues pour les caméras sphériques, c'est une extension de la géométrie deux et trois vues introduite par Torii et al. [21], où ils développent les contraintes bilinéaires et trilinéaires pour les caméras sphériques, mais où ils ne discutent pas des applications potentielles.

Considérons quatre caméras sphériques centrées en C_i (i=1..4) comme illustrées figure 1. Une ligne L de l'espace 3D est projetée sur les quatre images sphériques en grands cercles l_i définis par des normales notées n_i. L peut être définie vectoriellement par $L = X_0 + td$ où $L, X_0, d \in \mathbb{R}^3$ et $t \in \mathbb{R}$. $n_i \in \mathbb{R}^3$ étant les correspondances lignes/normales dans les quatre images sphériques.

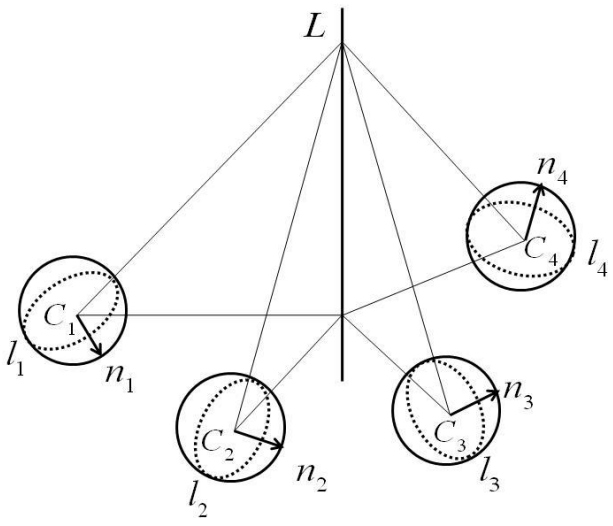


Figure 1: Géométrie quatre vues de caméras sphériques.

En supposant C_1 comme origine du repère monde, notons respectivement $[R_2|t_2]$, $[R_3|t_3]$ et $[R_4|t_4]$ les $[Rotation|translation]$ entre (C_1 et C_2), (C_1 et C_3) et (C_1 et C_4).

Comme une ligne 3D L est projetée sur des grands cercles l_i de normales respectives n_i dans C_i, nous avons les relations suivantes :

$$n_1^T L = 0 \tag{1}$$

$$n_2^T(R_2L + t_2) = 0 \quad (2)$$

$$n_3^T(R_3L + t_3) = 0 \quad (3)$$

$$n_4^T(R_4L + t_4) = 0 \quad (4)$$

2.1 Géométrie trois vues

Chaque triplet, c'est-à-dire chaque groupe de trois équations choisies parmi (1) à (4) donne une relation entre la ligne 3D L et ses normales correspondantes dans les trois vues. Par exemple, on peut obtenir à partir des vues 1,2,3 les équations (1),(2) et (3) que nous pouvons représenter par :

$$A\hat{L} = 0 \quad (5)$$

où

$$A = \begin{bmatrix} n_1^T & 0 \\ n_2^T R_2 & n_2^T t_2 \\ n_3^T R_3 & n_3^T t_3 \end{bmatrix} \quad \text{et} \quad \hat{L} = (L^T, 1)^T.$$

Comme ce système admet une solution non-nulle, la matrice A de taille 3×4 associée doit être de rang 2. Il existe donc une dépendance entre les trois lignes de A. Notons $A = (a_1^T, a_2^T, a_3^T)$, la relation linéaire entre ces trois lignes qui peut s'écrire $a_1 = \alpha a_2 + \beta a_3$. En remarquant que $a_{14} = 0$, on peut poser $\alpha = kt_3^T n_3$ et $\beta = -kt_2^T n_2$. Ainsi, nous obtenons les relations suivantes :

$$n_1^T = \alpha n_2^T R_2 + \beta n_3^T R_3$$

$$n_1 = \alpha R_2^T n_2 + \beta R_3^T n_3$$

$$n_1 = kt_3^T n_3 R_2^T n_2 - kt_2^T n_2 R_3^T n_3$$

Vues 1, 2 et 3:

$$R_2^T n_2 n_3^T t_3 - R_3^T n_3 n_2^T t_2 + \hat{k} n_1 = 0 \quad (6)$$

où $\hat{k} = -1/k$. Notons que k est nécessairement non-nul.
De manière identique, nous pouvons obtenir les relations suivantes avec les autres triplets :
Vues 1, 2 et 4:

$$R_2^T n_2 n_4^T t_4 - R_4^T n_4 n_2^T t_2 + \tilde{k} n_1 = 0 \quad (7)$$

Vues 1, 3 et 4:

$$R_3^T n_3 n_4^T t_4 - R_4^T n_4 n_3^T t_3 + \breve{k} n_1 = 0. \quad (8)$$

Chaque équation (6), (7) ou (8) relie les normales entre trois vues à partir des transformations $[R_i|t_i]$.

2.2 Géométrie quatre vues

Les équations (1) à (4) peuvent être réarrangées comme suit :

$$B\hat{L} = 0 \quad (9)$$

où

$$B = \begin{bmatrix} n_1^T & 0 \\ n_2^T R_2 & n_2^T t_2 \\ n_3^T R_3 & n_3^T t_3 \\ n_4^T R_4 & n_4^T t_4 \end{bmatrix} \quad \text{et} \quad \hat{L} = (L^T, 1)^T.$$

Une fois encore, comme le système (9) admet une solution non-nulle, la matrice B de taille 4×4 ne peut être inversible. Par conséquent, le déterminant de cette matrice est nul, ce qui se traduit par :

$$\begin{vmatrix} n_1^T \\ n_3^T R_3 \\ n_4^T R_4 \end{vmatrix} n_2^T t_2 - \begin{vmatrix} n_1^T \\ n_2^T R_2 \\ n_4^T R_4 \end{vmatrix} n_3^T t_3 + \begin{vmatrix} n_1^T \\ n_2^T R_2 \\ n_3^T R_3 \end{vmatrix} n_4^T t_4 = 0 \quad (10)$$

où $|.|$ représente le déterminant.
Ainsi, l'équation (10) relie les normales en correspondance entre les quatre vues à partir des transformations $[R_i|t_i]$.

3 Méthode d'estimation du mouvement

Configuration d'un banc stéréo: Considérons deux prises d'images obtenues par un banc stéréo : les deux premières images étant prises en C_1 et C_2 et les secondes en C_3 et C_4 comme illustré en figure 2. Avec cette configuration, les $[Rotation|translation]$ entre (C_3 and C_4) deviennent alors $[R_2|t_2]$. Ainsi, nous obtenons les contraintes suivantes :

$$\begin{cases} R_4 = R_2 R_3 \\ t_4 = t_2 + R_2 t_3 \end{cases} \quad (11)$$

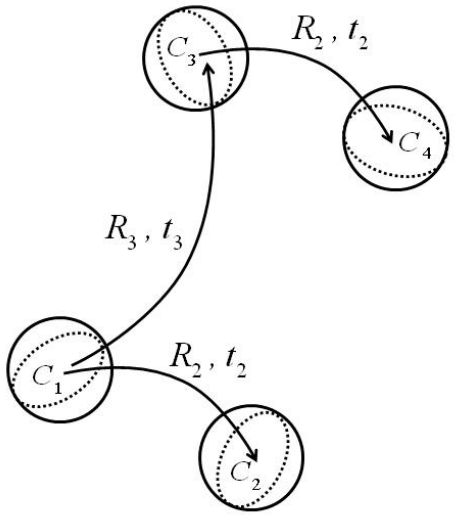


Figure 2: Banc stéréo sphérique

Ainsi, dans ce cas, seuls les mouvements $[R_2|t_2]$ et $[R_3|t_3]$ sont inconnus. Nous présentons dans la section suivante l'estimation de ces mouvements par découplage de la rotation et de la translation. C'est-à-dire, que nous estimons dans un premier temps les rotations R_2 et R_3 puis les translations t_2,t_3.

3.1 Rotation

La rotation entre les deux vues peut être calculée par la correspondance de deux points de fuites [4]. Ainsi nous détectons dans un premier temps les points de fuites dans les images sphériques à partir des faisceaux de droites parallèles présents dans la scène et nous calculons la rotation à partir de ces points de fuite mis en correspondance entre les images sphériques comme proposée par Horn dans [11].
La rotation R_2 est donc estimée comme suit:

$$\begin{cases} V_2 = R_2 V_1 \\ V_4 = R_2 V_3 \end{cases} \quad (12)$$

R_3 représente la rotation du banc stéréo entre deux positions arbitraires notées (C_1, C_2) et (C_3, C_4). R_3 peut donc être estimée à partir des points de fuites des vues 1 et 3 par $V_3 = R_3 V_1$. Néanmoins, nous pouvons aussi utiliser les points de fuite V_2 et V_4 des vues 2 et 4. Théoriquement, $V_1 = R_2^{-1} V_2$ et $V_3 = R_2^{-1} V_4$ mais en pratique, cette égalité peut ne pas être vérifiée à cause des erreurs d'estimation des V_i. C'est pourquoi, nous utilisons les points de fuite moyens afin d'estimer R_3 :

$$\frac{V_3 + R_2^{-1} V_4}{2} = R_3 \frac{V_1 + R_2^{-1} V_2}{2} \quad (13)$$

3.2 Estimation de la translation par géométrie trois vues

Il est possible d'estimer les translations t_2 et t_3 à partir (6), t_2 et t_4 à partir (7), t_3 et t_4 à partir (8). Ainsi, nous pouvons concaténer (6), (7) et (8) dans un système linéaire et à partir de la contrainte (11) estimer t_2, t_3 :

$$\begin{cases} R_2^T n_2 n_3^T t_3 - R_3^T n_3 n_2^T t_2 + \hat{k} n_1 = 0 \\ R_2^T n_2 n_4^T t_4 - R_4^T n_4 n_2^T t_2 + \tilde{k} n_1 = 0 \\ R_3^T n_3 n_4^T t_4 - R_4^T n_4 n_3^T t_3 + \breve{k} n_1 = 0 \end{cases} \Leftrightarrow MX = 0 \quad (14)$$

où

$$M = [M_L, M_R] \quad (15)$$

$$M_L = \begin{bmatrix} -R_3^T n_3 n_2^T & R_2^T n_2 n_3^T \\ R_2^T n_2 n_4^T - R_4^T n_4 n_2^T & R_2^T n_2 n_4^T R_2 \\ R_3^T n_3 n_4^T & R_3^T n_3 n_4^T R_2 - R_4^T n_4 n_3^T \end{bmatrix} \quad (16)$$

$$M_R = \mathtt{diag}(n_1, n_1, n_1) \quad (17)$$

$$X = (t_2^T, t_3^T, \hat{k}, \tilde{k}, \breve{k})^T \quad (18)$$

où $\hat{k}$, $\tilde{k}$ et $\breve{k}$ sont les scalaires associés aux triplets de vues respectives {1,2,3}, {1,2,4} et {1,3,4}.
Nous pouvons noter qu'un seul triplet peut suffire pour estimer t_2 et t_3. Cependant, comme les triplets de vues aboutissent à des équations indépendantes, nous utilisons tous les triplets de vues pour une meilleure robustesse de la méthode. En effet, au vue de la dernière colonne de la matrice M dans (15), il est impossible qu'un triplet dépende des autres.
Ainsi, à partir des correspondances lignes/normales $\{n_1, n_2, n_3, n_4\}$ dans les quatre vues sphériques, nous obtenons à partir de (14) un système linéaire en fonction des translations t_2 et t_3 et les trois scalaires. Chaque correspondance augmente la matrice M de 9 lignes et 3 colonnes et le vecteur d'inconnues X de 3 scalaires.
Avec N correspondances, nous obtenons le système suivant :

$$\hat{M}\hat{X} = 0 \quad (19)$$

où $\hat{M}$ est une matrice de taille $9N$x(6+$3N$) et $\hat{X}$ un vecteur de taille (6+$3N$) :

$$\hat{M} = [\hat{M}_L, \hat{M}_R] \quad (20)$$

où
$\hat{M}_L$ =

$$\begin{bmatrix} -R_3^T n_{31} n_{21}^T & R_2^T n_{21} n_{31}^T \\ R_2^T n_{21} n_{41}^T - R_4^T n_{41} n_{21}^T & R_2^T n_{21} n_{41}^T R_2 \\ R_3^T n_{31} n_{41}^T & R_3^T n_{31} n_{41}^T R_2 - R_4^T n_{41} n_{31}^T \\ \dots & \dots \\ \dots & \dots \\ -R_3^T n_{3N} n_{2N}^T & R_2^T n_{2N} n_{3N}^T \\ R_2^T n_{2N} n_{4N}^T - R_4^T n_{4N} n_{2N}^T & R_2^T n_{2N} n_{4N}^T R_2 \\ R_3^T n_{3N} n_{4N}^T & R_3^T n_{3N} n_{4N}^T R_2 - R_4^T n_{4N} n_{3N}^T \end{bmatrix} \quad (21)$$

$$\hat{M}_R = \mathtt{diag}(n_{11}, n_{11}, n_{11}, \dots n_{1N}, n_{1N}, n_{1N}) \quad (22)$$

$$\hat{X} = (t_2^T, t_3^T, \hat{k}_1, \tilde{k}_1, \breve{k}_1, \dots \hat{k}_N, \tilde{k}_N, \breve{k}_N)^T \quad (23)$$

où

- n_{ij} est la $j$ème correspondance dans la $i$ème vue (i=1..4, j=1..N)
- $\hat{k}_j$, $\tilde{k}_j$ et $\breve{k}_j$ sont respectivement les scalaires dans les trois triplets {1,2,3}, {1,2,4} et {1,3,4} pour la $j$ème correspondance

3.3 Estimation de la translation par géométrie quatre vues

L'équation (10) décrit la contrainte géométrique des correspondances des normales $\{n_1, n_2, n_3, n_4\}$ dans les quatre vues. N correspondances définissent donc un système linéaire par rapport aux translations t_2 et t_3 comme suit :

$$Q \begin{pmatrix} t_2 \\ t_3 \end{pmatrix} = 0 \quad (24)$$

où

$$Q = \begin{bmatrix} \Delta_{L1} & \Delta_{R1} \\ \Delta_{L2} & \Delta_{R2} \\ \dots & \dots \\ \Delta_{LN} & \Delta_{RN} \end{bmatrix} \quad (25)$$

$$\Delta_{Lj} = \begin{vmatrix} n_{1j}^T \\ n_{3j}^T R_3 \\ n_{4j}^T R_4 \end{vmatrix} n_{2j}^T + \begin{vmatrix} n_{1j}^T \\ n_{2j}^T R_2 \\ n_{3j}^T R_3 \end{vmatrix} n_{4j}^T \quad (26)$$

$$\Delta_{Rj} = \begin{vmatrix} n_{1j}^T \\ n_{2j}^T R_2 \\ n_{3j}^T R_3 \end{vmatrix} n_{4j}^T - \begin{vmatrix} n_{1j}^T \\ n_{2j}^T R_2 \\ n_{4j}^T R_4 \end{vmatrix} n_{3j}^T \quad (27)$$

où n_{ij} représente la $j$ème correspondance dans la $i$ème vue (i=1..4, j=1..N).

Par conséquent, à l'aide d'une décomposition en valeur singulière (SVD), nous pouvons résoudre le systèmes (19) ou (24) et ainsi obtenir les translations t_2 et t_3.
Sans considérer les effets du bruit dans la détection des normales dans les images ou dans l'estimation des rotations, nous avons analysé le rang des systèmes (19) et (24) générer à partir de $N \geq 2$ correspondances de lignes dans le tableau 1. Notons que $\hat{M}$ est une matrice de taille $9N$x(6+3N), ainsi trois correspondances suffisent pour estimer t_2 et t_3 à partir de trois triplets de vues de la section 3.2. La matrice Q est de taille Nx6, les translations peuvent donc être estimées dès que l'on a au moins cinq lignes mises en correspondance à partir du quadruplet de la section 3.3.

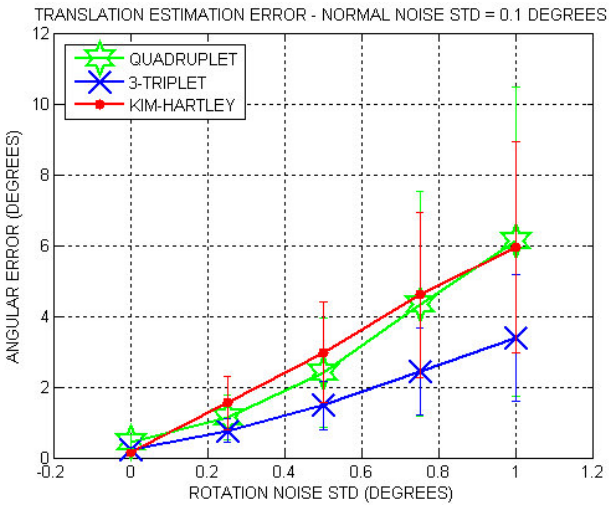
N	rang($\hat{M}$)	rang(Q)
2	10	min(N,5)
plus que 2	5+3N	

Table 1: Rang des systèmes linéaires.

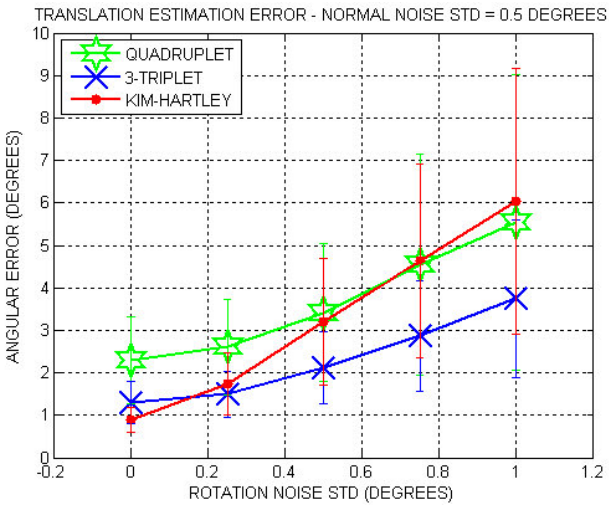
4 Résultats experimentaux

4.1 Données de synthèse

Nous avons créé dans un premier temps des lignes 3D autour de quatre caméras sphériques centrées en C_i (i=1..4). La distance moyenne entre deux positions du banc stéréo étant de 1000 mm et la distance moyenne entre les lignes 3D et les caméras est de 15000 mm. Ces lignes sont projetées sur les images sphériques sur des grands cercles. Les translations t_2, t_3 sont déterminées à partir de ces droites. Dans ce premier cas, on suppose parfaitement connues les rotations entre les caméras. Afin d'évaluer nos résultats, nous avons aussi implémenté la méthode d'estimation de la translation proposée par Kim et Hartley dans [12] où les translations des caméras sphériques sont estimées à partir des points.
Comme les normales de notre méthode et les points dans l'approche de Kim-Hartley se trouvent sur des sphères unitaires, ils peuvent être caractérisés par leurs angles d'élévation et d'azimuth. Pour vérifier la robustesse de notre méthode, nous avons bruité nos données par un bruit blanc gaussien d'écarts types 0.1 et 0.5 degrés sur les deux angles (azimuth/élévation) de chaque angle de chaque normale et de chaque point. Afin d'estimer l'influence du calcul de la précision de l'estimation de la rotation dans la méthode, nous avons perturbé les trois angles des rotations par des bruits blancs gaussiens d'écarts types de 0 à 1 degré. La figure 3 montre les erreurs moyennes des translations obtenues après 1000 essais.
On peut voir que l'utilisation des triplets de vues permet d'obtenir de meilleurs résultats qu'avec le quadruplet. Ceci s'explique par le fait que la matrice Q (24) est mal conditionnée. En effet, ces éléments sont tous proches de zéro car les Δ_{Lj} de (26) sont composés des relations trilinéaires des triplets (1, 2 et 3) et (1, 3 and 4).



(a)



(b)

Figure 3: Erreur d'estimation de la translation par les approches quadruplet, 3-triplet et Kim-Hartley [12].

Comparée à la méthode de Kim-Hartley [12], notre approche permet d'obtenir de meilleures estimations des translations. De plus, il faut noter que l'emploi des droites est plus favorable que celui des points dans le cas d'images sphériques provenant de capteurs hétérogènes que ce soit pour la détection que pour la mise en correspondance. Il est nécessaire d'ajouter également que bruiter la normale d'un grand cercle a plus de conséquence sur l'estimation du déplacement que le bruitage d'un point.

4.2 Images réelles

Expérience 1 : Nous avons déplacé un banc stéréo hybride muni d'une caméra perspective et d'une caméra fisheye et estimer le déplacement en deux prises d'images à partir de la méthode 3-triplet décrite dans ce papier. Nous montrons figure 4 un exemple d'images prises par ce banc ainsi que l'estimation du déplacement. C_1C_2, la baseline du banc stéréo est de 45 cm. C_1C_3 représente le déplacement du banc entre les deux prises d'images. Le ratio $(C_1C_2 : C_1C_3)$ est utilisé pour comparer l'estimation par rapport à la vérité terrain. En bas à gauche de la figure 4, la vérité terrain est de $(C_1C_2 : C_1C_3 = 45cm : 200cm = 1.00 : 4.44)$ et nous retrouvons $(C_1C_2 : C_1C_3 = 1.00 : 4.59)$. En bas à droite de la figure 4, la vérité terrain est de $(C_1C_2 : C_1C_3 = 45cm : 700cm = 1.00 : 15.56)$ et nous retrouvons $(C_1C_2 : C_1C_3 = 1.00 : 15.72)$.

Expérience 2 : Nous estimons dans cette expérience le mouvement d'un banc stéréo composé de deux caméras para-catadioptriques. Le déplacement effectué est tracé en rouge sur la figure 5 où les images consécutives du banc ont été prises à une distance constante de 68 cm. La figure 6 permet de comparer notre méthode avec l'algorithme proposé par Kim-Hartley [12].

5 Conclusion

Nous avons présenté dans ce papier une méthode d'estimation de déplacement d'un banc stéréoscopique composé de caméras hybrides à point de vue unique. Cette méthode utilise exclusivement les droites comme primitives. Nous avons validé notre méthode sur des données de synthèses et réelles en la comparant avec une méthode d'estimation basée sur les points développée par Kim-Hartley [12]. Nous avons remarqué que lors de l'estimation du déplacement à partir des quatre vues, la méthode basée sur les 3-triplets est préférable à une approche basée sur le quadruplet. A l'aide du modèle de projection sphérique unifié, cette approche peut être appliquée à tous types de caméras à point de vue unique. De plus, nous avons utilisé uniquement les primitives lignes qui présentent un avantage par rapport aux méthodes basées points pour des caméras hybrides car il est en effet très difficile de détecter les mêmes points d'intérêt sur des caméras hétérogènes. De plus, l'utilisation d'un seul type de primitive pour estimer à la fois les rotations et les translations permet d'obtenir une méthode rapide et donc utilisable en robotique mobile.

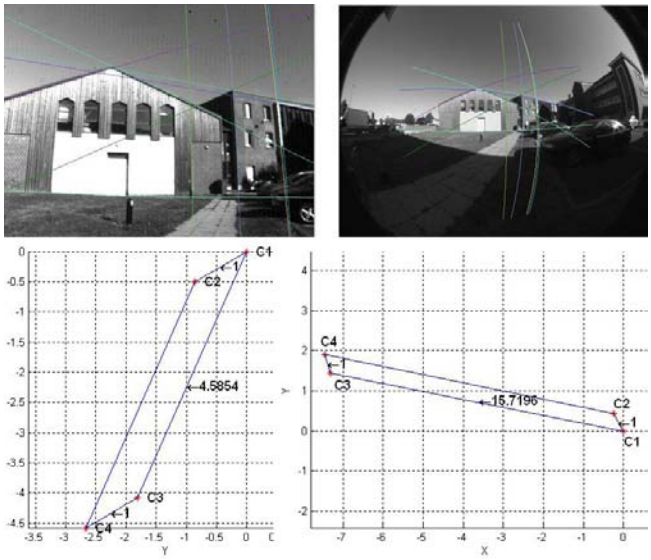


Figure 4: Expérience 1 - Estimation du déplacement d'un banc stéréo perspectif-fisheye : (en haut) exemple d'images avec la détection des lignes, (en bas) estimation de deux déplacements du banc à partir de l'approche 3-triplet.

6 Remerciements

Ce travail est soutenu par le projet PHC STAR NOVA (19038YH) et le Projet Région Picardie ALTO (Axe thématique de rattachement : Intermodalité, Energie, Transport).

References

[1] M. E. Antone and S. J. Teller. Scalable extrinsic calibration of omnidirectional image networks. In *International Journal of Computer Vision (IJCV)*, vol. 49, pp. 143-174, 2002.

[2] J. P. Barreto and H. Araujo. Issues on the geometry of central catadioptric image formation. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, vol. 2, pp. 422-427, 2001.

[3] H. Bay, V. Ferrari and L.J. Van Gool. Wide-baseline stereo matching with line segments. In *CVPR*, pp. 329-336, 2005.

[4] J.C. Bazin, C. Demonceaux, P. Vasseur and I.S. Kweon. Motion estimation by decoupling rotation and translation in catadioptric vision. In *Computer Vision and Image Understanding (CVIU)*, 2009.

[5] A. I. Comport, E. Malis and P. Rives. Accurate Quadrifocal Tracking for Robust 3D Visual Odometry. In *International Conference on Robotics and Automation (ICRA)*, pp. 40-45, 2007.

[6] P. David, D. Dementhon, R. Duraiswami and H. Samet. Simultaneous pose and correspondence deter-

Figure 5: Expérience 2 - Exemple d'images de la séquence

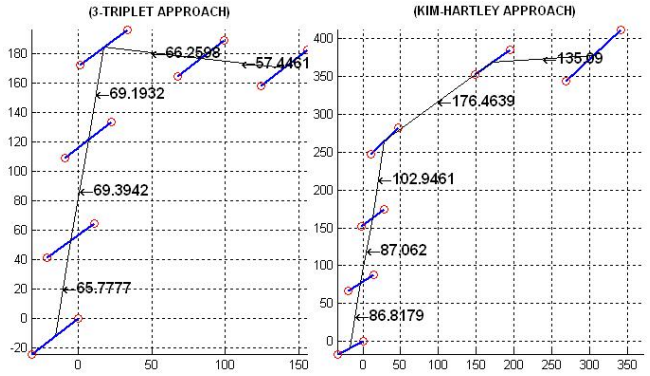


Figure 6: Expérience 2 - Estimation du déplacement du banc stéréo paracatadioptrique par les approches 3-triplet (gauche) et Kim-Harley [12] (droite)

mination using line features. In *CVPR*, vol. 2, pp. 424, 2003.

[7] C. Geyer and K. Daniilidis. A unifying theory for central panoramic systems and practical implications. In *Proc. of the European Conference on Computer Vision (ECCV)*, pp. 445-461, 2000.

[8] J. Gluckman, S.K. Nayar and K.J. Thoresz. Real-Time omnidirectional and Panoramic Stereo. In *Proceedings of the 1998 DARPA Image Understanding Workshop*, Monterey, California, pp. 209-303, 1998.

[9] P. Gros, O. Bournez and E. Boyer. Using local planar geometric invariants to match and model images of line segments. In *CVIU*, vol. 69, no. 2, pp. 135-155, 1998.

[10] R. Hartley and A. Zisserman. *Multiple view geometry in computer vision*. Cambridge University, 2nd edition, 2003.

[11] B. K. P. Horn. Closed-form solution of absolute orientation using unit quaternions. In *Journal of the Optical Society of America. A*, vol. 4, no. 4, pp. 629-642, 1987.

[12] J.H. Kim and R. Hartley. Translation estimation from omnidirectional images. In *Digital Image Computing: Techniques and Applications (DICTA)*, pp. 22, 2005.

[13] T. Lemaire, C. Berger, I. K. Jung and S. Lacroix. Vision-Based SLAM: Stereo and Monocular Approaches. In *IJCV*, vol. 74, no. 3, pp. 343-364, 2007.

[14] S. Li. Real-Time Spherical Stereo. In *International Conference on Pattern Recognition (ICPR)*, pp. 1046-1049, 2006.

[15] A. Mallet, S. Lacroix and L. Gallo. Position estimation in outdoor environments using pixel tracking and stereovision. In *ICRA*, pp. 3519-3524, 2000.

[16] R. Mandelbaum, G. Salgian and H. Sawhney. Correlation-based estimation of ego-motion and structure from motion and stereo. In *Proc. of IEEE International Conf. on Computer Vision (ICCV)*, vol. 1, pp. 544, 1999.

[17] C. Mei. Laser-augmented omnidirectional vision for 3D localisation and mapping. PhD Thesis, 2007.

[18] M. Ollis, H. Herman and S. Singh. Analysis and Design of Panoramic Stereo Vision Using Equi-Angular Pixel Cameras. In *Technical report CMU-RI-TR-99-04*, Robotics Institute, Carnegie Mellon University, 1999.

[19] C. Olson, L. Matthies, M. Schoppers and M. Maimone. Robust stereo ego-motion for long distance navigation. In *CVPR*, vol. 2, pp. 453-458, 2000.

[20] P. Sturm. Mixing Catadioptric and Perspective Cameras. In *Proc. of IEEE Workshop on Omnidirectional Vision (OMNIVIS02)*.

[21] A. Torii, A. Imiya and N. Ohnishi. Two- and three-view geometry for spherical cameras. In *Proc. of IEEE Workshop on Omnidirectional Vision (OMNIVIS05)*.

[22] J. Weng, P. Cohen and N. Rebibo. Motion and structure estimation from stereo image sequences. In *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, 8(3):362-382, 1992.

[23] X.H. Ying and Z.Y. Hu. Can We Consider Central Catadioptric Cameras and Fisheye Cameras within a Unified Imaging Model. In *ECCV*, vol. I, pp. 442-455, 2004.

[24] Z. Zhang and O. Faugeras. Estimation of displacements from two 3D frames obtained from stereo. In *PAMI*, 14(12):1141-1156, 1992.