

Matrice semi fondamentale pour système de stéréoscopie hybride

Cylindrical Calibration for Hybrid Epipolar Configurations

Laurent Smadja, Ryad Benosman et Jean Devars

Laboratoire des Instruments et Systèmes d'Ile-de-France, Université Paris VI, BP 164, Tour 22-12 2^{ème} étage
4 Place Jussieu, 75252 Paris Cedex 05, France, email{laurent.smadja@lis.jussieu.fr,{benosman|devars}@ccr.jussieu.fr}

Manuscrit reçu le 24 janvier 2005

Résumé et mots clés

Cet article présente le calibrage d'un capteur cylindrique omnidirectionnel, qui fournit des panoramas Haute Résolution par la mise en rotation d'une caméra CCD linéaire.

L'algorithme utilise les segments de droite pour primitives, dont les projections sont des composées de sinusoides, et est constitué de deux étapes. Un algorithme de détection et d'identification de l'image de droites est proposé, selon une méthode proche de la transformée de Hough. De plus, l'utilisation des coordonnées de Plücker permet la linéarisation de la projection cylindrique.

Les résultats extraits de la première étape sont utilisés pour l'évaluation du premier paramètre intrinsèque. Le second est quant à lui déterminé par l'intermédiaire de la minimisation d'un critère linéaire dans l'espace d'accumulation.

Calibrage non linéaire, Imagerie Panoramique, Coordonnées de Plücker, Transformée de Hough.

Abstract and key words

This paper deals with the calibration of a cylindrical omnidirectional imaging system, based on a rotating 2048 pixels linear camera which provides high definition panoramas.

The two-step algorithm relies on line segment projections as calibration features, which are sinusoidal curves. A cylindrical image line detection algorithm is proposed, based on the dual principle of the Hough transform. Moreover, the use of Plücker coordinates introduces some new characteristics in the calibration process. This kind of formalism allows a linearization of the cylindrical projection, which is non-linear in the usual way.

Results obtained from this first step are used to evaluate one of the intrinsic parameters, the other one being determined by a linear criterion minimization in the dual space, *ie* the sine magnitude space.

Non-Linear Calibration, Omnidirectional Imaging System, Plücker, Coordinates, Hough Transform.

1. Introduction

L'étalonnage constitue l'une des étapes majeures de tout algorithme de vision par ordinateur, car elle fournit la relation entre les points image et les directions vers les points 3D auxquels ils sont associés. Cette étape a été très largement étudiée [Hartley and Zissermann, 2000, Faugeras, 1993, Tsai and Lenz, 1989], et il existe aujourd'hui différentes techniques d'étalonnage adaptées aux capteurs classiques, s'appuyant sur plusieurs primitives, tels que les coins, les segments de droite, les cercles (ou ellipses), voire les surfaces de révolution [Won *et al.*, 2003].

La vision omnidirectionnelle, depuis ses débuts il y a plus de quinze ans [Cao and Oh, 1986, Yagi and Kawato, 1990, Ishiguro *et al.*, 1992], a récemment connu un grand développement [Omn, Benosman and Kang, 2001]. On peut principalement distinguer deux courants : la voie la plus explorée s'appuie sur l'association de caméras perspectives ou affines et de miroirs de courbure connue. Ce type de capteurs, appelés *catadioptriques*, peut être modélisé par une projection centrale [Geyer and Daniilidis, 2000, Nayar, 1997, Baker and Nayar, 1999] si la courbure du miroir est paraboloidale ou hyperboloidale. Les autres miroirs, d'une manière générale, ne satisfont certes pas la contrainte du point focal unique, mais peuvent présenter d'autres avantages, comme par exemple une résolution constante selon certaines directions particulières [Gaspar *et al.*, 2002, Swaminathan *et al.*, 2004]. Les systèmes à projection centrale semblent aujourd'hui relativement bien maîtrisés, et sont de plus en plus usités dans le cadre de la navigation de robots mobiles, ou la télé-surveillance. Cependant, ces capteurs sont généralement dotés d'une résolution assez pauvre et variable. De plus, la fabrication du miroir doit être particulièrement précise, ainsi que le positionnement du capteur CCD par rapport à celui-ci, ce qui reste assez difficile à réaliser pratiquement.

La seconde approche pour produire des images panoramiques s'appuie sur la mise en rotation de caméras perspectives, afin de créer des projections cylindriques [Benosman, 1997, Shum *et al.*, 1999, Smadja, 2003]. Ce type de capteurs ne peut évidemment pas être appliqué à un procédé temps réel, de par le temps d'acquisition, mais fournit des images de bien meilleure résolution. On peut citer ici les travaux de Peleg ou Shum [Peleg and Ben Ezra, 1999, Shum and Szeliski, 1999], qui sélectionnent deux colonnes de pixels d'une caméra mise en rotation pour produire en une révolution un couple de panoramas stéréo. La distance entre les deux points de vue, dite la *base*, reste cependant trop faible pour obtenir des disparités exploitables pour les objets lointains. Benosman [Benosman *et al.*, 1996] utilise deux caméras linéaires alignées verticalement et mises en rotation, la reconstruction est ensuite effectuée à l'aide de l'ensemble des procédés propres à la stéréovision, la mise en correspondance étant grandement facilitée de par la configuration épipolaire particulièrement simple. On peut cependant regretter que l'étalonnage utilisé n'exploite pas la géométrie cylindrique de l'image et se restreigne au format linéaire de la caméra.

D'une manière générale, les capteurs omnidirectionnels ne peuvent être calibrés d'une manière classique, du fait de la non-linéarité de leur surface image. Les segments de droite semblent constituer, par contre, une primitive nettement plus adaptée que les points pour cette étape. Par exemple, Geyer et Daniilidis [Geyer and Daniilidis, 2002] utilisent un jeu de droites (parallèles ou non) afin de déterminer les paramètres d'une caméra para-catadioptrique (association d'un miroir de forme paraboloidale et d'une caméra orthographique). De la même manière, Huang [Huang, 2000] prouve que les algorithmes d'étalonnage à base de lignes peuvent avantageusement être employés dans le cadre de panoramas cylindriques.

Le capteur développé se base sur toutes ces observations : l'objectif est de générer des cartes de disparité haute définition d'environnements panoramiques. L'approche séquentielle a donc été préférée puisqu'elle est à même de fournir des images dont la résolution est assez fine. D'autre part, le système ne contient qu'une seule caméra, ce qui implique la nécessité de prendre plusieurs images pour toute étude stéréo mais permet de choisir la base du couple en fonction des dimensions de la scène à reconstruire.

L'article est organisé comme suit. La section 2 présente le modèle géométrique du capteur cylindrique développé. Nous présentons également les bases de l'étude formelle du système, et précisons les choix quant à la primitive à exploiter. Les coordonnées de Plücker permettent une linéarisation de la projection de segments de droite sur le cylindre image, et ainsi d'obtenir un modèle exploitable dans le cadre de l'étalonnage.

La section 3 présente une méthode proche de la transformée de Hough, permettant la détection et l'identification de l'image de segments de droite dans un panorama cylindrique. Cet algorithme donne également une estimation de l'un des paramètres intrinsèques par deux méthodes différentes, et fournit les données nécessaires à l'étalonnage complet, dont les résultats sont exposés section 4.

L'association d'images cylindrique et classique conduit à des configurations épipolaires inhabituelles. La géométrie épipolaire exprime la relation, généralement linéaire, entre plusieurs images prises selon des points de vue différents. La nature non-linéaire de la projection cylindrique ne peut pas être intégrée dans la matrice essentielle, de manière à obtenir une contrainte stéréoscopique non-calibrée. Nous introduisons alors dans la section 5 une matrice *semi-fondamentale* ${}_pF_c$ (où p correspond à perspectif et c à cylindrique), qui traduit la géométrie épipolaire dans ce type de configurations hybrides et peut être déterminée à partir de plus de 8 appariements dans un couple perspectif/cylindrique.

2. Modèle Cylindrique et étalonnage

Le développement de la projection d'un point sur le capteur cylindrique montre rapidement que cette primitive n'est pas bien adaptée à l'étalonnage. L'équation obtenue est fortement non-linéaire et ne peut que difficilement être exploitée pour la détermination des paramètres relatifs à la formation de l'image. L'utilisation de primitives différentes, telles que les segments de droites, par l'intermédiaire du formalisme de Plücker, permet une linéarisation du modèle de projection.

2.1. Architecture mécanique

Le capteur est basé sur l'utilisation d'une caméra linéaire, mise en rotation sur un axe parallèle à la barrette elle-même par un moteur à courant continu. Celui-ci envoie régulièrement, typiquement tous les dixièmes de degré, un signal à la caméra déclenchant l'acquisition d'une colonne [figure 1]. En concaténant les colonnes acquises au cours d'une rotation complète, on obtient une image panoramique cylindrique RGB.

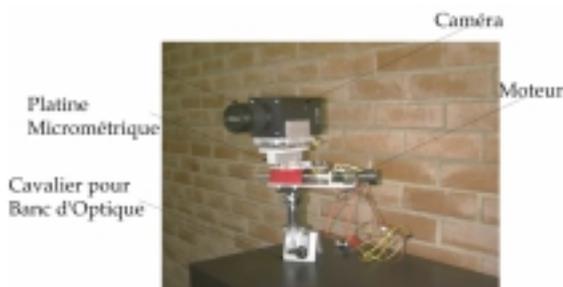


Figure 1. Montage global: la caméra est montée sur un banc d'optique. Une platine micrométrique permet un positionnement précis.

Pour obtenir une image à point focal unique, il faut que le centre optique de la barrette CCD soit aligné avec l'axe de rotation du moteur. En effet, il est important de vérifier cette contrainte sous peine d'obtenir des images à caustique circulaire, ce qui en modifie largement la modélisation. Une méthode précise et simple dans sa mise en œuvre permet de positionner le centre de la caméra. Pour plus de détails, le lecteur pourra se reporter à [Smadja *et al.*, 2001].

Les vues omnidirectionnelles obtenues avec ce procédé possèdent une excellente définition: une image du couple stéréo a pour dimensions $2048 \times 3600 \times 3$ pixels [figure 2]

2.2. Image d'un Point

De par la géométrie du capteur, le plan image peut être modélisé par une surface cylindrique centrée autour du centre optique de la barrette CCD. La primitive élémentaire en traitement d'image étant le point, il apparaît logique de commencer la modélisation mathématique par l'expression des coordonnées de la projection de points sur l'image cylindrique.

Soit le repère *Cylindre* R_C , repère dont l'origine est le centre optique de la barrette CCD, et dont l'axe Z correspond à l'axe de révolution du cylindre. Un point P de coordonnées $[X_c, Y_c, Z_c]^T$ dans R_C , est projeté en p sur le cylindre, le «plan» image [figure 3] via:

$$p \begin{cases} x = f \frac{X_c}{Z_c} \\ y = f \frac{Y_c}{Z_c} \\ z = f \frac{Z_c}{R} \end{cases} \quad (1)$$

où f est la distance focale équivalente du capteur, c'est-à-dire le rayon du cylindre et $R = \sqrt{X_c^2 + Y_c^2}$.



Figure 2. Un exemple d'image réalisée avec le capteur haute définition.

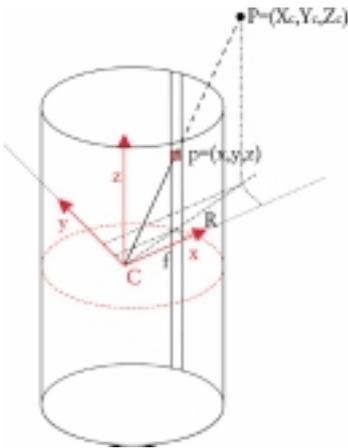


Figure 3. Projection de $P(X_c, Y_c, Z_c)$ en $p(x, y, z)$.

Les coordonnées du point image m sont quant à elles directement reliées aux coordonnées cylindriques de p (r, θ, z), par la transformation :

$$m \begin{cases} u = K_u \theta = K_u \arctan \left(\frac{Y_c}{X_c} \right) \\ v = v_0 - K_v z = v_0 - K_v f \frac{Z_c}{R} \end{cases} \quad (2)$$

où K_u et K_v sont les pas de discrétisation de l'image et v_0 la projection du centre optique du capteur sur la barrette CCD. Le paramètre K_v correspond plus précisément à l'inverse de la dimension verticale d'un pixel, alors que K_u est le pas discrétisation angulaire. Ces trois paramètres sont les *paramètres intrinsèques*.

Pour être complet, le modèle caméra doit exprimer la relation entre des points 3D exprimés dans un repère propre, le *repère Monde* R_M , et leurs coordonnées images. La transformation rigide (R, \mathbf{t}) qui lie ces deux repères est paramétrée par les angles d'Euler et \mathbf{t} , la translation entre les deux centres optiques : il s'agit des paramètres *extrinsèques*.

Il paraît cependant difficile d'intégrer cette dernière transformation dans l'équation (2). Il est ainsi impossible de modéliser *linéairement* la projection de points de R_M , en coordonnées pixelliques sur le cylindre image : d'autres primitives sont donc à considérer.

2.3. Image d'un Segment de Droite

Reprenons les coordonnées du point P , exprimées dans R_C . Si ce point est situé sur un segment, la droite support de ce segment définit avec l'origine du repère caméra un plan, dit *plan de vue*. Les coordonnées de P vérifient alors :

$$N_1 X_c + N_2 Y_c + N_3 Z_c = 0 \implies \frac{Z_c}{R} = -\frac{N_1}{N_3} \frac{X_c}{R} - \frac{N_2}{N_3} \frac{Y_c}{R} \quad (3)$$

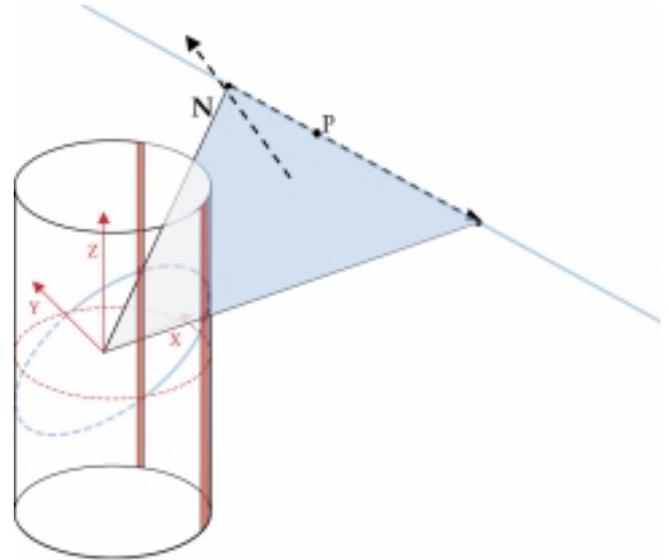


Figure 4. Projection d'une droite sur le cylindre : son image dépend du plan de vue correspondant N .

Combiner les équations (2) et (3), en remarquant que $\frac{X_c}{R} = \cos \theta$ et $\frac{Y_c}{R} = \sin \theta$, conduit à :

$$m \begin{cases} u = K_u \theta \\ v = v_0 + K_v f \frac{N_1}{N_3} \cos \left(\frac{u}{K_u} \right) + K_v f \frac{N_2}{N_3} \sin \left(\frac{u}{K_u} \right) \end{cases} \quad (4)$$

Si l'on considère particulièrement la seconde des deux dernières équations, on déduit que l'image d'un segment de droite par le capteur cylindrique est une *composée de sinusoides* dont les amplitudes sont données par les rapports des coordonnées du plan image relatif à ce segment [figure 4].

L'expression (4) est pour le moment restreinte au repère R_C et doit être exprimée dans R_M , par l'intermédiaire des paramètres extrinsèques.

2.4. Formalisme de Plücker

Les vecteurs de Plücker [Plücker, 1865, Shoemake, 1998] semblent être une représentation simple des droites de \mathbb{R}^3 permettant force facilités de calcul. Nous présentons ici quelques définitions et développements relatifs à ce formalisme.

2.4.1. Définitions

1. Tout vecteur de six coordonnées homogènes (soit deux vecteurs à trois coordonnées) $\{\mathbf{M}, \mathbf{N}\}$ tel que $\mathbf{M} \cdot \mathbf{N} = 0$ correspond à une droite de \mathbb{R}^3 .
2. Le plan (quatre coordonnées homogènes) défini par la droite $L = \{\mathbf{M}, \mathbf{N}\}$ passant par le point $P = (\mathbf{P}, w)$ est donné par $(\mathbf{M} \times \mathbf{P} - w \cdot \mathbf{N}, \mathbf{N} \cdot \mathbf{P})$

3. Deux droites $L_1 = \{\mathbf{M}_1, \mathbf{N}_1\}$ et $L_2 = \{\mathbf{M}_2, \mathbf{N}_2\}$ sont coplanaires si et seulement si $\mathbf{M}_1 \cdot \mathbf{N}_2 + \mathbf{M}_2 \cdot \mathbf{N}_1 = 0$

4. La droite passant par P et Q est donnée par $L = \{P - Q, P \times Q\}$

Ce dernier point est particulièrement intéressant, il permet de donner aux vecteurs \mathbf{M} et \mathbf{N} une signification physique. Le vecteur \mathbf{M} est évidemment le vecteur directeur de la droite considérée. Le vecteur \mathbf{N} , quant à lui, est le produit vectoriel de deux vecteurs respectivement formés par deux points de la droite et l'origine du repère, *i.e.*, le plan de vue. Le formalisme de Plücker semble donc particulièrement adapté à notre capteur car il encode presque directement l'image d'un segment de droite par le capteur cylindrique.

2.4.2. Changement de Repère

On introduit à présent les paramètres dits *extrinsèques*, qui traduisent le déplacement rigide (rotation R , translation \mathbf{t}) entre deux repères. Comment un vecteur $L = \{\mathbf{M}, \mathbf{N}\}$ représentant une droite de \mathbb{R}^3 est-il transformé lors d'une isométrie?

- Rotation pure R : Les deux vecteurs \mathbf{M} et \mathbf{N} subissent la rotation comme un simple produit à gauche:

$$L' = \{R\mathbf{M}, R\mathbf{N}\} \quad (5)$$

- Translation pure \mathbf{t} : Si le vecteur \mathbf{M} ne subit aucune transformation quand le repère est simplement translaté, le vecteur \mathbf{N} est lui, par contre, largement modifié. En effet \mathbf{N} représente la normale au plan formé par la droite et le nouveau centre du repère. La deuxième définition nous donne les coordonnées de ce plan, dans le cas où $P = (\mathbf{t}, \mathbf{1})$: $(\mathbf{M} \times \mathbf{t} - \mathbf{N}, \mathbf{N} \cdot \mathbf{t})$. On obtient alors, après une translation pure:

$$L' = \{\mathbf{M}, \mathbf{M} \times \mathbf{t} - \mathbf{N}\} \quad (6)$$

- Isométrie: En combinant les deux déplacements, on obtient la modification d'un vecteur de Plücker après un changement de repère:

$$L' = \{R\mathbf{M}, R(\mathbf{M} \times \mathbf{t} - \mathbf{N})\} \quad (7)$$

2.5. Modèle Complet

Considérons une droite de R_M , de coordonnées $L = [\mathbf{M}', \mathbf{N}']^T$. De par la formule (7), la normale au plan de vue \mathbf{N} peut être exprimée linéairement dans R_C :

$$\mathbf{N}_{R_C} = \begin{pmatrix} N_1 \\ N_2 \\ N_3 \end{pmatrix}_{R_C} \sim [-RT_{\times}|R] \begin{bmatrix} \mathbf{M}' \\ \mathbf{N}' \end{bmatrix}_{R_M} \quad (8)$$

où T_{\times} est la matrice anti-symétrique correspondant au produit vectoriel par le vecteur \mathbf{T} et la matrice $[RT_{\times}|R]$ a pour dimension 3×6 .

Rappelons qu'une sinusoïde image (u, v) a pour équation:

$$v = v_0 + A \cos \alpha u + B \sin \alpha u \quad \text{où } \alpha = \frac{1}{K_u}, \quad (9)$$

on peut identifier les amplitudes A et B des composantes des sinusoïdes aux rap-

ports $K_v f \frac{N_1}{N_3}$ et $K_v f \frac{N_2}{N_3}$ de l'équation (4):

$$\begin{pmatrix} A \\ B \\ 1 \end{pmatrix} = K_v f \cdot \begin{pmatrix} \frac{N_1}{N_3} \\ \frac{N_2}{N_3} \\ 1 \end{pmatrix}_{R_C} \sim K_v f \cdot \begin{pmatrix} N_1 \\ N_2 \\ N_3 \end{pmatrix}_{R_C} \quad (10)$$

En regroupant les équations (8) et (10), on aboutit à un modèle regroupant les paramètres extrinsèques et l'un des paramètres intrinsèques, $K_v f$. Ce modèle lie les vecteurs de Plücker $L = [\mathbf{M}', \mathbf{N}']^T_{R_M}$ de droites exprimées dans un repère quelconque aux amplitudes (A, B) des sinusoïdes images qui leur sont associées:

$$\begin{pmatrix} A \\ B \\ 1 \end{pmatrix} \sim K_v f \cdot [-RT_{\times}|R] \begin{bmatrix} \mathbf{M}' \\ \mathbf{N}' \end{bmatrix}_{R_M} \quad (11)$$

Cette équation sera très utile pour l'étape d'étalonnage et implique deux des trois paramètres intrinsèques:

- K_u , qui correspond à la « période » de révolution du capteur. Sachant en effet qu'une image panoramique couvrant 360 possède K_u colonnes régulièrement réparties, ce paramètre est directement lié au format de l'image, il n'est donc pas à déterminer.

- v_0 , la projection du centre optique sur la barrette CCD.

- $K_v f$, la focale, *i.e.* le rayon du cylindre image, exprimé en pixels.

Les 6 paramètres extrinsèques, 3 pour la rotation et 3 pour la translation sont également à déterminer, ce qui porte à 8 le nombre d'inconnues. Chaque droite image détectée fournit deux équations, reliant les vecteurs de Plücker aux amplitudes des sinusoïdes:

$$\begin{cases} A = K_v f \frac{N_1}{N_3} \\ B = K_v f \frac{N_2}{N_3} \end{cases} \quad (12)$$

où $\mathbf{N} = [N_1, N_2, N_3] = -[RT_{\times}|R] L$.

Quatre correspondances sinusoïdes / vecteurs de Plücker sont donc nécessaires dans le cas minimal, pour déterminer l'ensemble des paramètres restants.

3. Détection de Droites

La détection de droites est un sujet relativement connu et éprouvé pour des images perspectives classiques. Nous présentons ici une approche dérivée de la transformée de Hough permettant leur détection sur panoramas cylindriques.

3.1. Transformée de Hough classique: Rappels

L'algorithme de Hough [Hou62] s'appuie sur la formulation normale des droites exprimée par Duda et Hart [Duda and Hart, 1972]. Une droite peut en effet aussi bien être représentée par le couple (ρ_0, θ_0) que par le «traditionnel» (a, b) :

$$y = \frac{-1}{\tan\theta_0}x + \frac{\rho_0}{\sin\theta_0} \implies \rho_0 = x \cos \theta_0 + y \sin \theta_0 \quad (13)$$

Un point de l'espace image (x, y) est ainsi associé à une sinusoïde dans l'espace de Hough (ρ, θ) . L'accumulation des sinusoïdes correspondant à l'ensemble des points de contours permet d'obtenir, *via* la position des maxima locaux, les différents segments de droites de l'image d'origine.



En tenant compte du fait que toute image panoramique peut se reprojeter sur une surface cylindrique, nous présentons un algorithme s'appuyant sur la théorie de Hough, permettant la détection de sinusoïdes images.

3.2. Transformée de Hough Duale

Nous avons en effet prouvé dans la section 2.3 que toute droite de l'espace 3D se projette sur le capteur cylindrique selon une composée de sinusoïdes (*cf.* équation 4). Rappelons que, dans le cas classique, la transformée de Hough transpose des droites image en sinusoïdes duales. Il existe donc un parallèle évident, si l'on omet le décalage vertical v_0 , entre la transformée classique et celle que l'on veut développer. Le problème est inverse à celui résolu par Hough, et l'on peut ainsi, en appliquant l'algorithme inverse, construire des cartes d'accumulation de droites dont les maxima locaux fourniront les amplitudes des sinusoïdes recherchées [figure 5]. On s'affranchit pour le moment du décalage v_0 par une simple translation, nous reviendrons sur ce point par la suite, et l'on applique l'algorithme suivant :

Pour chaque point de contour (u, v) de l'image d'origine :

{ Calculer les valeurs $a = \frac{-1}{\tan \frac{u}{K_u}}$ et $b = \frac{v - offset_v}{\sin \frac{u}{K_u}}$ (eq. 13).

{ En faisant varier x entre X_{min} et X_{max} , tracer la droite

$$y = ax + b \text{ dans l'espace dual de Hough.}$$

Extraire les maxima locaux de la carte d'accumulation, leurs coordonnées x et y correspondent respectivement à

$$A = K_v f \frac{n_1}{n_3} \text{ et } B = K_v f \frac{n_2}{n_3}$$

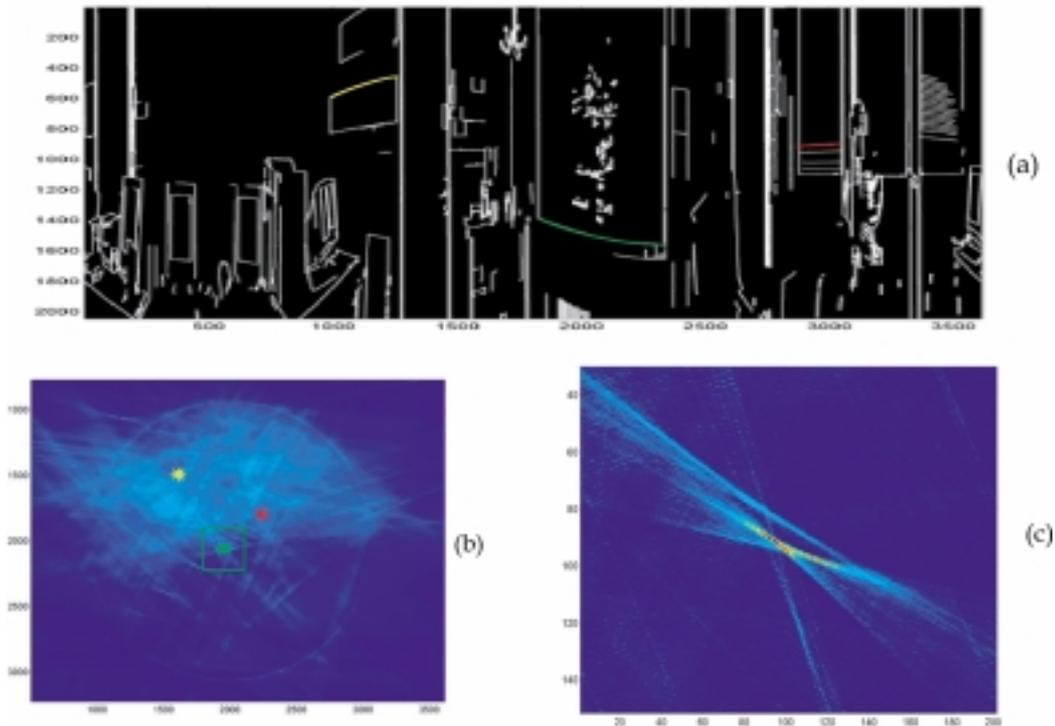


Figure 5. (a) Une image panoramique des contours. (b) La carte de Hough dual associée à cette image. (c) Grossissement autour d'un des maxima locaux.

3.3. Erreur Relative à un Biais sur l'Estimation de v_0

Les sinusoides images associées aux droites de \mathbb{R}^3 ne sont pas centrées en 0, car nos images cylindriques ont, comme pour la plupart des caméras classiques, leur origine dans le coin supérieur gauche. Cependant, l'équation (4) indique v_0 comme le décalage vertical des courbes. Un simple changement de repère, une translation de $offset_v$ dans la direction verticale de l'image nous renvoie une image centrée sur cette valeur. Il existe cependant un biais pour une valeur de v_0 différente de la vérité terrain, qui induit une altération des résultats attendus et qui dépend de l'erreur commise. Il s'agit alors de définir une méthode robuste propre à déterminer la valeur optimale de v_0 ; deux approches ont été abordées, la première basée sur la carte d'accumulation elle-même, la seconde par un retour dans le plan image. Ces deux algorithmes s'appuient sur plusieurs cartes duales construites autour de décalages différents.

3.3.1. Influence de v_0 sur la Carte des Votes : Critère dual

Concrètement, une erreur sur la valeur de v_0 ($offset_v$) n'agit que sur $\rho = v - v_0$ et ne change donc que l'ordonnée à l'origine des différentes droites duales. Quand on commet une erreur δv_0 , les coordonnées duales (A, B) correspondant aux accumulations de sinusoides sont modifiées telles que :

$$\begin{cases} A' = A + \delta v_0 \left(\frac{\sin \theta_1 - \sin \theta_2}{\sin \theta_1 \cos \theta_2 - \sin \theta_2 \cos \theta_1} \right) \\ B' = B - \delta v_0 \left(\frac{1 + \cos \theta_2}{\sin \theta_1} \frac{\sin \theta_2 - \sin \theta_1}{\sin \theta_2 \cos \theta_1 - \sin \theta_1 \cos \theta_2} \right) \end{cases} \quad (14)$$

L'équation (14) montre clairement que les valeurs des intersections sont déplacées selon un terme additif proportionnel à l'erreur commise δv_0 (sur v_0) mais dépendant également de la

position angulaire u_i des votants considérés; il y a donc une dispersion autour des accumulations optimales [figure 6]. On doit ainsi pouvoir observer une diminution de la hauteur des pics d'autant plus importante que le biais sur v_0 est grand. On peut alors estimer une valeur optimale du paramètre intrinsèque quand les maxima correspondant aux sinusoides images passent par leur plus haut maximum. La dispersion est en effet minimale quand le décalage sur lequel a été construit la carte correspond exactement à la vérité terrain. Nous avons utilisé l'image des contours d'une mire d'étalonnage spécifique; les cartes d'accumulation construites autour de différentes valeurs de v_0 présentent des maxima locaux, dont la hauteur moyenne constitue notre premier critère.

3.3.2. Identification des Votants : Élaboration d'un Critère Image

La seconde méthode utilisée s'appuie sur un retour dans le plan image et passe par une identification des votants. Il s'agit en effet de pouvoir identifier sur la carte duale les points de contour ayant voté pour une sinusoides particulière, associée à l'un des maxima locaux.

On obtient, en calculant la somme quadratique des distances images entre les points de contour et leur sinusoides, une erreur associée à une carte centrée sur une valeur de v_0 . Cette étape est réalisée à l'aide d'un deuxième « passage » sur l'ensemble des cartes issues de l'algorithme précédent :

Pour chaque point de contour de l'image d'origine :

{Recalculer les valeurs a et b et évaluer la distance de la droite obtenue à l'ensemble des maxima locaux dans l'espace dual. On rappelle que la distance d'un point $[x_0, y_0, 1]$ à une droite $[a, -1, b]$ est donnée par la formule :

$$d = \frac{ax_0 - y_0 + b}{\sqrt{a^2 + 1}}$$

S

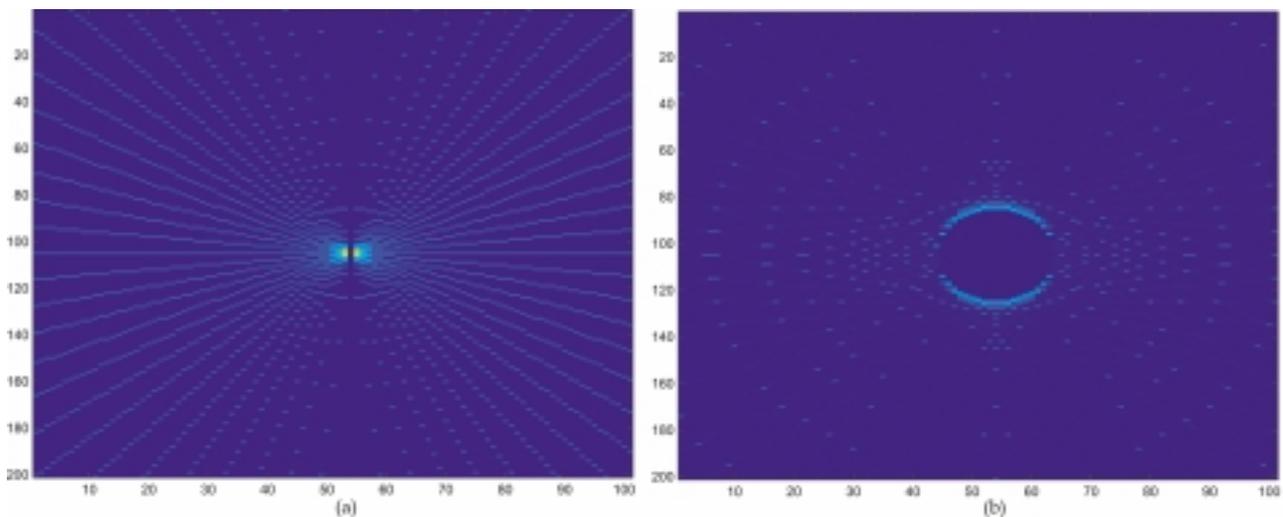


Figure 6. (a) En simulation, l'accumulation est parfaite pour une valeur correcte de v_0 .
(b) Il existe une dispersion des votes proportionnel au biais sur v_0 .

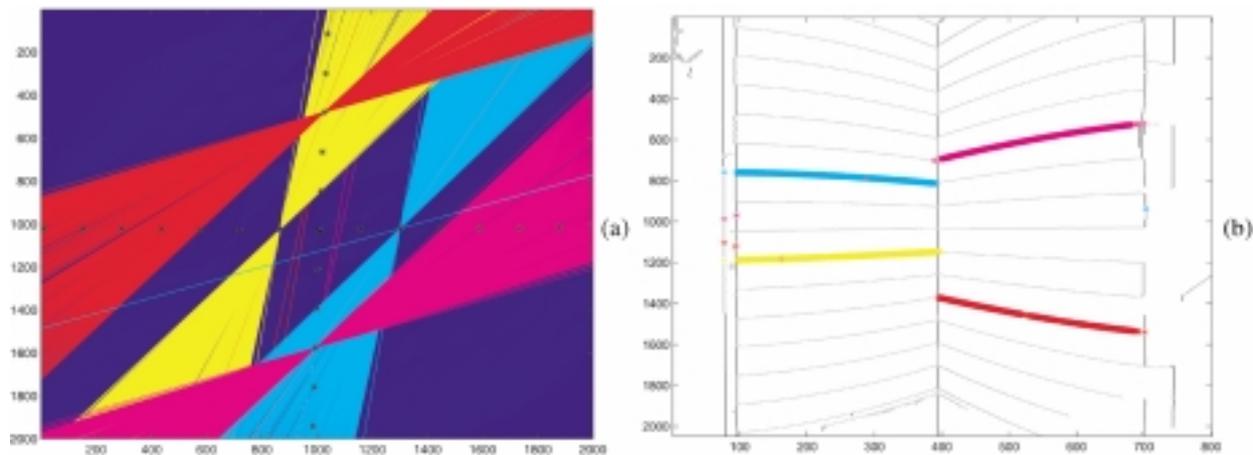


Figure 7. (a) Quelques maxima locaux et les droites duales ayant voté pour eux dans l'espace dual. (b) Les sinusôides associées sur l'image panoramique.

{Associer ce point de contour à une sinusôide image si cette distance est inférieure à un seuil. Dans la pratique, ce seuil est choisi empiriquement et dépend fortement du pas de discrétisation et de la taille totale de la carte duale.

{Calculer la distance image quadratique entre le point de contour et la sinusôide pour laquelle il a voté.

{Mettre à jour le critère.



On obtient une erreur image pour tous les points de contour, dont la moyenne correspond à l'erreur associée à chacune des valeurs de v_0 . C'est la valeur renvoyant l'erreur image la plus faible qui est finalement choisie.

4. Résultats

4.1. Détection et Détermination de v_0

Nous avons utilisé une mire constituée d'ensemble de lignes disposées sur deux plans perpendiculaires, à l'image des mires de

points généralement employées. On calcule une image des contours de cette mire par l'intermédiaire d'un détecteur de Canny, ce qui fournit les données nécessaires à l'élaboration des cartes d'accumulation autour de v_0 différents. La barette CCD employée possède 2048 pixels, le v_0 « idéal » serait placé au milieu de celle-ci; nous avons donc choisi un intervalle compris entre 980 et 1200, ce qui nous donne un total de 221 cartes duales.

4.1.1 Critère Image

L'identification des votants nécessaires au calcul du critère image est présentée figure 7. Cette étape permet ainsi, par une différence image quadratique entre chaque point de contour et la sinusôide pour laquelle il a voté, d'estimer le critère image associé à une carte d'accumulation, et donc à une valeur de v_0 . La carte présentant l'erreur normalisée par rapport au nombre de votants correspond alors à la meilleure estimation. Ce critère image renvoie une valeur optimale de 1067 comme paramètre intrinsèque v_0 , qui correspond à une erreur moyenne de 0.4059 pixel par points de contour [figure 8].

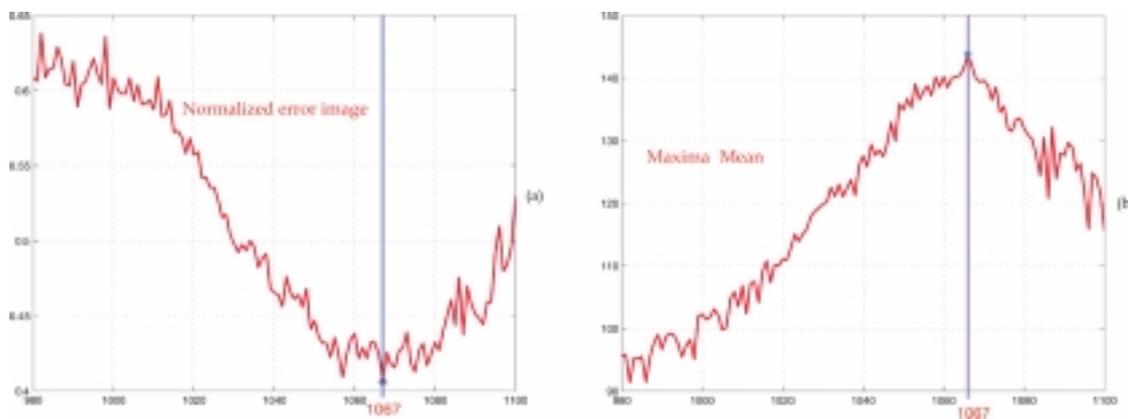


Figure 8. (a) Évolution de l'erreur image en fonction de v_0 : l'erreur minimale est atteinte à $v_0 = 1067$. (b) Évolution de la moyenne de la hauteur des maxima locaux en fonction de v_0 : la plus haute moyenne renvoie $v_0 = 1067$.

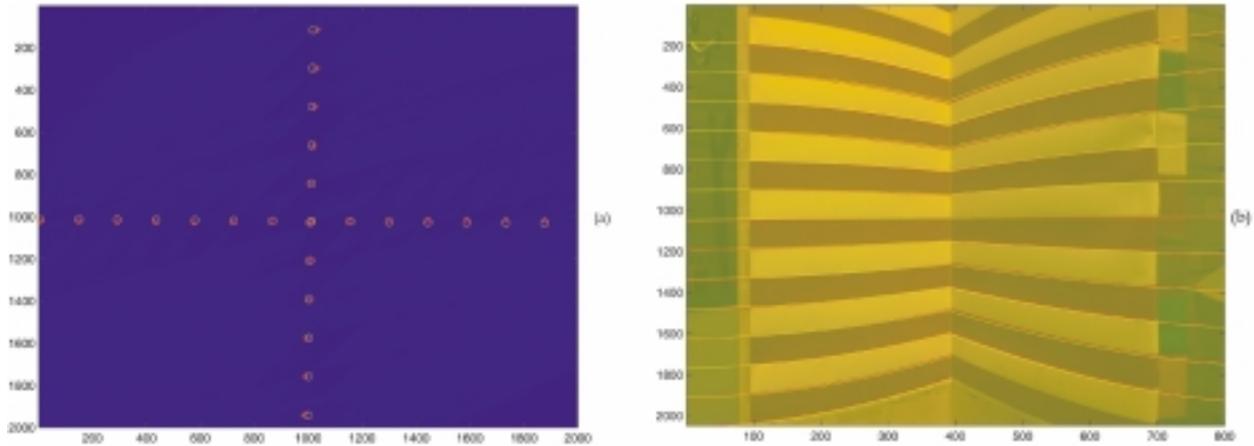


Figure 9. Après minimisation de la distance dans l'espace dual : (a) Les coordonnées obtenues. (b) Les droites de la mire après l'étape d'étalonnage : En rouge, les droites détectées, en jaune, les droites reconstruites.

4.1.2. Critère Dual

Les évolutions des hauteurs de chaque maximum local montrent en effet une augmentation de la hauteur des pics quand v_0 se rapproche de 1 065. Il est clair que l'ensemble des sinusoides, *i.e.* les maxima de chaque carte, n'indiquent pas la même valeur comme optimale, mais la moyenne des pics, ainsi que sa médiane, renvoient une valeur proche de **1067**, associée à une hauteur moyenne de pics de **140** votants, ce qui corrobore les résultats issus de la minimisation du critère image [figure 8].

La valeur finalement retenue pour le premier paramètre intrinsèque v_0 est donc de **1067**, résultat qui vérifie les deux critères présentés. Il correspond en effet à la meilleure moyenne des hauteurs de pics, ce qui atteste de la plus petite dispersion des votants. Il minimise de plus le second critère, c'est-à-dire la plus petite erreur image normalisée. Cette valeur, ainsi que la position des maxima locaux dans l'espace dual, vont permettre d'effectuer la seconde phase du procédé d'étalonnage.

4.2.2. Étalonnage Complet

Les amplitudes de sinusoides, extraites pour $v_0 = 1067$, sont à présent utilisées pour la détermination des paramètres restants, par l'intermédiaire de l'équation (11).

Le procédé de minimisation s'appuie sur l'algorithme de *Levenberg-Marquardt*, dans \mathbb{R}^7 . Le vecteur de paramètres obtenu avec les données correspondant à $v_0 = 1067$ vaut :

$$\left\{ \begin{array}{lll} \xi = -0.0077 & \theta = 3.1374 & \varphi = 4.7017 \\ tx = 546.98 & ty = 693.75 & tz = 212.22 \\ & K_v f = 3274.7 & \end{array} \right\} \quad (15)$$

avec une erreur de 2.9157. Cette mesure est sans unité, c'est une distance euclidienne dans l'espace des amplitudes de sinusoides, elle a donc peu de signification physique.

Les positions reconstruites des maxima locaux dans la carte d'accumulation, ainsi que les sinusoides associées sont illus-

trées sur la figure 9.

Si les paramètres extrinsèques de rotation et de translation nous importent peu dans cette étape, le paramètre $K_v f$ est quant à lui très important pour la suite de l'étude. La focale, exprimée en pixels, est en effet le second paramètre intrinsèque du capteur cylindrique, tel que nous l'avons présenté dans la section 2.3.

5. Configuration Stéréo Hybride

Depuis l'article de Longuet-Higgins [Longuet-Higgins, 1981] introduisant la matrice *essentielle* E , les algorithmes de stéréovision s'appuient sur une contrainte géométrique pour faciliter la phase de mise en correspondance. Ce concept a depuis été adapté aux capteurs non-linéaires, et particulièrement avec plusieurs types de capteurs panoramiques [Bunschoten and Kröse, 2001, Cauchois *et al.*, 2003, Chang and Hebert, 2000].

Appliquée sur des images non calibrées, cette relation est formalisée par l'intermédiaire de la matrice *fondamentale*, nettement plus utilisée car elle traduit une contrainte relative aux coordonnées images des points à apparier. Le cas non calibré a ainsi été beaucoup plus étudié [Faugeras, 1993, Hartley and Zissermann, 2000, Zhang, 1998]. D'autre part, la matrice fondamentale peut être généralisée à plusieurs types de capteurs et à un plus grand nombre de vues, via l'utilisation de tenseurs. Sturm [Sturm, 2002] présente les relations inhérentes aux vues multiples, pour une configuration hybride faite de caméras classiques et de capteurs para-catadioptriques, par l'intermédiaire de tenseurs tri ou quadri-focaux.

En suivant la même approche, nous introduisons la matrice *semi fondamentale* ${}_p F_c$ (avec p pour perspectif et c pour cylindrique) issue de l'association d'un capteur cylindrique calibré et d'une caméra classique non calibrée. Cette matrice encode la contrainte

de coplanarité, inhérente à tout système de stéréovision, et peut être déterminée, à un facteur près, à partir de 8 correspondances de points dans le cas minimal.

5.1. Présentation Formelle

Considérons deux images prises par un capteur stéréo composé de deux caméras perspectives classiques, représentées par le modèle sténopé, un point P de la scène réelle et ses deux projections p_1 et p_2 . Dans la suite de ce paragraphe, les points image, exprimés en pixels, seront dénotés par m_i , alors que les points « métriques » seront, eux, représentés par p_i .

La géométrie épipolaire traduit simplement la coplanarité des 2 centres optiques O_1, O_2 et de ce point P , les vecteurs O_1p_1, O_2p_2 et O_1O_2 définissant le plan épipolaire.

Le point image p_2 apparié à p_1 appartient donc obligatoirement à l'intersection du second plan image avec le plan épipolaire. Les courbes paramétrées obtenues dépendent fortement de la géométrie du capteur: cela va des simples droites dans le cas classique, aux cercles dans le cas de capteurs sphériques et paraobloïdes ou d'ellipses dans le cas cylindrique [figure 10]. Nous avons en effet prouvé dans la section 2.3. que la projection de droites sur une surface cylindrique s'exprimait par une composée de sinusôïdes, qui correspond à une ellipse « déroulée ». Cette contrainte s'exprime par l'intermédiaire de la matrice essentielle E [Longuet-Higgins, 19891]:

$$p_1^T E p_2 = 0 \tag{16}$$

Considérons à présent que l'image 1 est une image perspective classique et l'image 2 un panorama cylindrique. La position métrique p_1 est linéairement liée à m_1 , exprimée cette fois en coordonnées images *via*:

$$m_1 = K p_1 \Rightarrow p_1^T = m_1^T K^{-T} \tag{17}$$

où K est la matrice triangulaire supérieure de la caméra perspective 1. En combinant (17) et (16), on aboutit à:

$$m_1^T K^{-T} E p_2 = 0 \Rightarrow m_1^T {}_pF_c p_2 = 0 \tag{18}$$

avec:

$${}_pF_c = K^{-T} E = K^{-T} T_x R \tag{19}$$

On obtient ainsi une matrice essentielle *semi-calibrée*, ${}_pF_c$, qui encode la contrainte épipolaire pour un système stéréo associant un capteur classique non calibré à une caméra panoramique calibrée.

5.2. Détermination de ${}_pF_c$

${}_pF_c$ est une matrice 3×3 homogène, donc définie à un facteur près. Chaque point apparié du couple stéréo fournit une seule contrainte; 8 paires sont donc suffisantes, dans le cas minimal, pour déterminer ${}_pF_c$. Nous avons utilisé le classique *8-point algorithm* [Hartley, 1995], brièvement présenté ici.

L'équation (18) peut se réécrire d'une manière linéaire selon les éléments de ${}_pF_c$. En effet, si $m_1 = [u_1, v_1, 1]^T$, $p_2 = [X_2, Y_2, Z_2]^T$ et ${}_pF_c$ est décomposé en un vecteur à neuf composantes, on obtient:

$$[u_1 X_2, u_1 Y_2, u_1 Z_2, v_1 X_2, v_1 Y_2, v_1 Z_2, X_2, Y_2, Z_2] \begin{bmatrix} f_{11} \\ f_{12} \\ \vdots \\ f_{33} \end{bmatrix} = 0 \tag{20}$$

En concaténant n_{points} contraintes ($n_{points} > 8$), exprimées par l'équation (20), dans une matrice A de dimensions $9 \times n_{points}$, on aboutit à un problème de minimisation bien connu: $Ag = 0$. La solution issue de l'utilisation de la décomposition en valeurs propres [Zhang, 1998] induit une normalisation de g , ce qui

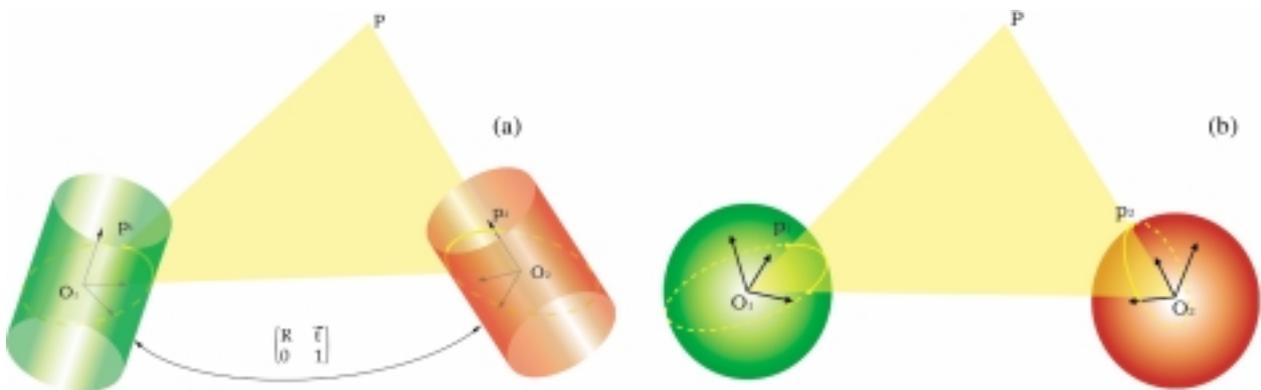


Figure 10. Quel que soit le modèle de projection, la courbe épipolaire correspond à l'intersection du plan épipolaire et de la surface image: (a) le cas cylindrique. (b) le cas sphérique.

écarte le cas trivial $g = 0$. Cette normalisation n'autorise cependant, dans le cas classique, la détermination de la translation qu'à un facteur près. Dans notre cas, elle influence également les valeurs de T_x et K . Cette approche, la *Direct Linear Transform* (DLT) fournit ainsi très facilement une première estimation pour ${}_pF_c$.

Par la suite, on peut raffiner ce premier résultat par l'utilisation de la contrainte de rang 2. De fait, ${}_pF_c$, tout comme E ou F , n'est pas une matrice de rang plein, car elle inclut la matrice antisymétrique T_x , qui est naturellement singulière. Tsai a démontré [Tsai and Huang, 1994] que si ${}_pF_c = UDV^T$, où $D = \text{diag}(d_1, d_2, d_3)$, est la décomposition en valeurs singulières de ${}_pF_c$, alors $\widehat{{}_pF_c} = U\widehat{D}V^T$, avec $\widehat{D} = \text{diag}(d_1, d_2, 0)$, minimise la norme de Frobenius de $\|{}_pF_c - \widehat{{}_pF_c}\|$ contraint par $\det_p \widehat{{}_pF_c} = 0$.

Ces deux étapes ont été utilisées pour déterminer ${}_pF_c$ pour un couple perspectif/cylindrique. On peut par ailleurs imposer une normalisation des données proche de celle préconisée par Hartley [Hartley, 1995], afin de mieux conditionner la matrice A . Les résultats obtenus sont présentés dans la section suivante.

5.3. Résultats

Plusieurs images perspectives classiques peuvent être associées à un simple panorama cylindrique. Ces combinaisons produisent autant de matrices ${}_pF_c$, selon la position de la caméra classique, se déplaçant dans l'environnement. Nous présentons ici une image classique, associée à un seul panorama cylindrique, mais l'on peut envisager une collection d'images planes, c'est-à-dire plusieurs matrices ${}_pF_c$ différentes.

Les points utilisés dans cette étape ont été sélectionnés manuellement, et aucune normalisation n'a été appliquée. La figure [11] présente ces points ainsi que les courbes épipolaires obtenues avec la solution aux valeurs singulières.

La solution obtenue à l'aide du DLT fournit des courbes épipolaires correctes, car celles-ci passent bien par les points correspondant aux points cliqués sur l'image perspective ([figure 11] en jaune sur l'image de droite). Cependant, on voit que l'intersection des différentes courbes, l'épipole, n'est pas unique.

Imposer la contrainte de rang 2 fournit de meilleurs résultats, les courbes épipolaires convergent alors vers un épipole unique

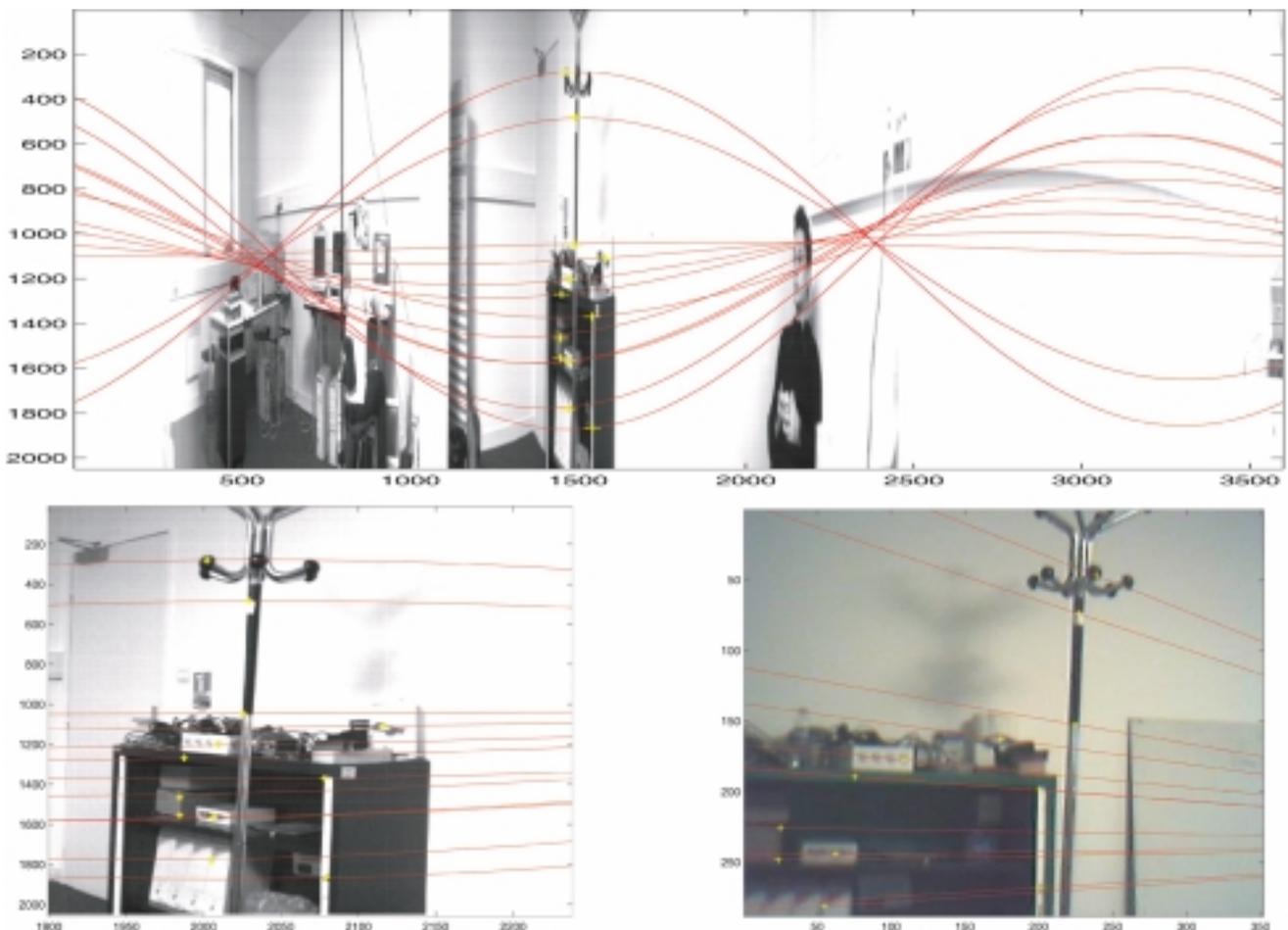


Figure 11. Les courbes épipolaires après une première estimation par DLT. En haut, l'image cylindrique. En bas à gauche, un détail du panorama

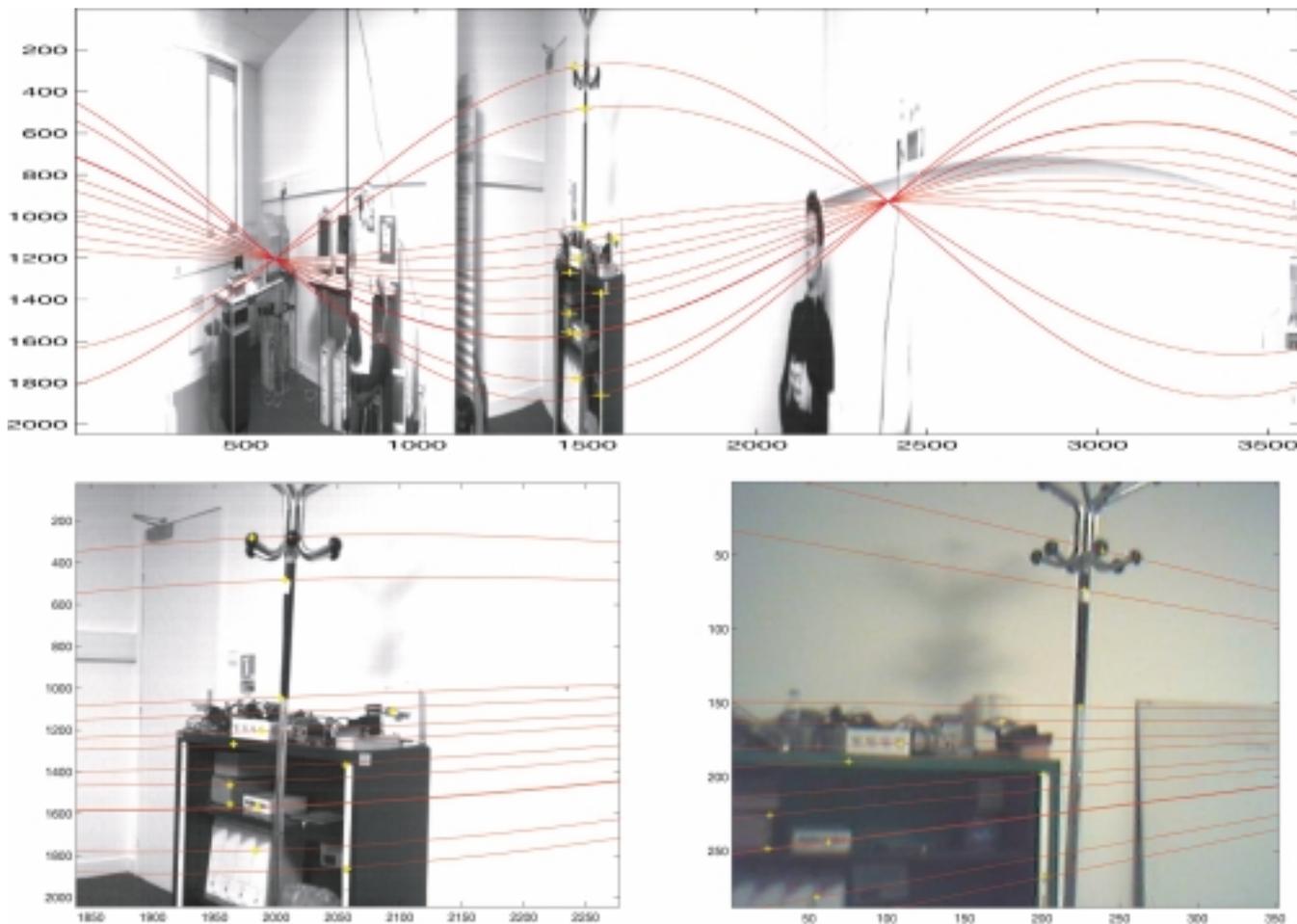


Figure 12. Après l'utilisation de la contrainte de rang 2: Les courbes épipolaires convergent vers un épipole unique.

[figure 12]. Cette contrainte assure en effet l'unicité de la direction de translation sans pour autant modifier la précision des courbes.

6. Conclusion

Cet article présente une formalisation globale de la projection relative à un capteur omnidirectionnel haute définition, qui produit des panoramas cylindriques de dimensions $2048 \times 3600 \times 3$.

Les algorithmes classiques d'étalonnage ne peuvent pas être appliqués à ce type de capteur, du fait de la non-linéarité de la surface de projection. Les droites de \mathbb{R}^3 constituent alors une primitive plus adéquate; leurs images par le capteur cylindrique sont des combinaisons de sinusoides, dont les amplitudes

sont liéairement reliées aux vecteurs de Plücker. L'utilisation de ce formalisme permet en effet la linéarisation de l'ensemble du modèle.

Un algorithme de détection de sinusoides images est ainsi présenté, afin de fournir une estimation de l'un des intrinsèques v_0 , selon deux approches différentes, ainsi que les données nécessaires à l'étalonnage complet. Cette méthode possède l'avantage de pouvoir s'adapter à plusieurs types de capteurs catadioptriques, dont les images distordues sont fréquemment « dépliées » sur des surfaces cylindriques.

L'association du capteur cylindrique et d'une caméra classique engendre des configurations épipolaires inédites. La matrice semi-fondamentale ${}_pF_c$ permet alors d'exprimer cette contrainte dans le cas général calibré/non-calibré. Nous présentons les résultats obtenus pour la détermination de cette matrice, et les différentes courbes épipolaires obtenues dans un contexte alliant caméra panoramique et complexe multi-caméras.

Références

- [Omn] *Proceedings of the IEEE Workshop on Omnidirectional Vision (OMNIVIS'00 05)*.
- [Baker and Nayar] S. BAKER, and S. NAYAR, A theory of single-viewpoint catadioptric image formation. *International Journal of Computer Vision*, pages 175-196, (1999).
- [Benosman] R. BENOSMAN, *Reconstruction d'environnements perçus par un capteur panoramique de stéréovision et détermination des modèles CAO*. PhD thesis, Université Pierre et Marie Curie, (1997).
- [Benosman and Kang] R. BENOSMAN and S. KANG, *Panoramic Vision: Theory, Sensors and Applications*. Springer Verlag, (2001).
- [Benosman et al.] R. BENOSMAN, T. MANIÈRE and J. DEVARs, Multi-directional stereovision sensor, calibration and scenes reconstruction. *13th International Conference on Pattern Recognition*, A:161-165, (1996).
- [Bunschoten and Kröse] R. BUNSCHOTEN and B. KRÖSE, Range estimation from a pair of omnidirectional images. *Proc. IEEE Int. Conf. on Robotics and Automation*, pages 1174-1179, (2001).
- [Cao and Oh] Z.L. CAO, S. J OH and E.L. HALL, Dynamic omnidirectional vision for mobile robots. *Journal of Robotics Systems*, 3(1):5-17, (1986).
- [Cauchois et al.] C. CAUCHOIS, E. BRASSART E, L. DELAHOUCHE and A. CLERENTIN, 3d localization with conical vision. *Proc. IEEE Workshop on Omnidirectional Vision OMNIVIS'03*, (2003).
- [Chang and Hebert] P. CHANG and M. HEBERT, Omnidirectional structure from motion. *IEEE Workshop on Omnidirectional Vision (OMNIVIS)*, (2000).
- [Duda and Hart] R. O. DUDA and P.E. HART, Use of the hough transform to detect lines and curves in pictures. *Communications of the ACM*, 15:11-15, (1972).
- [Faugeras] O. FAUGERAS, *Three-Dimensional Computer Vision*. MIT Press. (1993)
- [Gaspar et al.] J. GASPAS, C. DECCO, J. J. OKAMOTO and J. SANTOS-VICTOR, Constant resolution omnidirectional cameras. *Proc. of IEEE Workshop on Omnidirectional Vision Omnivis' 02*, pages 27-34, (2002).
- [Geyer and Daniilidis] C. GEYER and K. DANIILIDIS, A unifying theory for central panoramic systems and practical applications. *Proc of ECCV*, pages 445-461, (2000).
- [Geyer and Daniilidis] C. GEYER and K. DANIILIDIS, Para-catadioptric calibration. *IEEE Trans. Pattern Analysis and Machine Intelligence*, (2002).
- [Hartley] R. HARTLEY, In defence of 8-point algorithm. *Proc. 5th Int. Conference on Computer Vision*,, pages 1064-1070, (1995).
- [Hartley and Zissermann] R. I. HARTLEY and A. ZISSERMANN, *Multiple view geometry in computer vision*. Cambridge University Press, (2000).
- [Huang.] F. HUANG, *Epipolar geometry and camera calibration of cylindrical panoramas*. PhD thesis, Auckland University, (2000).
- [Ishiguro et al.] I. ISHIGURO, M. YAMAMOTO and S. TSUJI, Omnidirectional stereo. *IEEE Trans. on PAMI*, 14:257-262, (1992).
- [Longuet-Higgins] H. LONGUET-HIGGINS, A computer algorithm for reconstructing a scene from two projections. *Nature*, 293:133-135, (1981).
- [Nayar] S. K. NAYAR, Catadioptric omnidirectional camera. *Proc. of IEEE CVPR*, pages 482-488, (1997).
- [Peleg and Ben Ezra] S. PELEG and M. BEN EZRA, Stereo panoramic with a single camera. *proc. of IEEE CVPR*, 1:395-401, (1999).
- [Plücker] J. PLÜCKER, On a new geometry of space. *Philosophical Transactions of the Royal Society of London*, pages 725-791, (1865).
- [Shoemake] K. SHOEMAKE, Plucker coordinate tutorial. *Ray-Tracing News vol 11*, (1998).
- [Shum et al.] H. Y. SHUM, A. KALAI and S. SEITZ, Omnivergent stereo. *Proc. Seventh International Conference on Computer Vision (ICCV'99)*, (1999).
- [Shum and Szeliski] H. Y. SHUM and R. SZELISKI, Stereo reconstruction from multiperspective cameras. *Proc. of IEEE ICCV*, 1:14-21, (1999).
- [Smadja] L. SMADJA, *Génération d'environnements 3D denses à partir d'images panoramiques cylindriques*. PhD thesis, UPMC, (2003).
- [Smadja et al.] L. SMADJA, R. BENOSMAN and J. DEVARs, Determining epipolar constraint on cylindrical images and using it for 3d reconstruction. *Proc. ICAR*, (2001).
- [Sturm] P. STURM, Mixing catadioptric and perspective cameras. *Proceedings of the Workshop on Omnidirectional Vision*, (2002).
- [Swaminathan et al.] R. SWAMINATHAN, S. NAYAR and M. GROSSBERG, Designing mirrors for catadioptric systems that minimize image errors. *IEEE Workshop on Omnidirectional Vision Omnivis'04*, (2004).
- [Tsai and Lenz] R. TSAI and R. LENZ, A new technique for fully autonomous and efficient 3d robotics hand/eye calibration. *IEEE Trans Robotics Automation*, 5:345-358, (1989).
- [Tsai and Huang] R. Y. TSAI and T. S. HUANG, Uniqueness and estimation of three dimensional motion parameters of rigid objects with curved surfaces. *IEEE Trans. on PAMI*, 6:13-27, (1994).
- [Won et al.] K. Y. WON, P. R. S. MENDOÇA and R. CIPOLLA, Camera calibration from surfaces of revolution. *IEEE Trans. on PAMI*, 25(2):147-161, (2003).
- [Yagi and Kawato] Y. YAGI and S. KAWATO, Panorama scene analysis with conic projection. In *Proceedings IEEE/RSJ International Workshop on Intelligent Robots and Systems '90*, pages 181-187, Tsuchiura, (1990).
- [Zhang] Z. ZHANG, Determining the epipolar geometry and its uncertainty: A review. *International Journal of Computer Vision*, 27(2):161-198, (1998).



Laurent **Smadja**

Laurent Smadja a soutenu une thèse en informatique à l'université Pierre et Marie Curie (Paris6) en 2003, et actuellement ATER dans l'équipe Vision. Pendant son doctorat réalisé dans le Laboratoire des Instruments et Systèmes d'Ile de France (LISIF), il a développé un système d'imagerie panoramique et formalisé l'ensemble des procédés propres à la stéréoscopie adaptée à l'imagerie non-linéaire. Ses thématiques de recherche concernent la géométrie des vues multiples, et les systèmes stéréoscopiques hybrides.



Ryad **Benosman**

Ryad Benosman a soutenu une thèse en robotique à l'université Pierre et Marie Curie (Paris6) en 1997. Il est actuellement maître de Conférences au Laboratoire des Instruments et Systèmes d'Ile de France, dans l'équipe Vision. Ses thématiques de recherches sont la vision Omnidirectionnelle, la vision embarquée (robotique mobile). Il travaille actuellement les réseaux dynamiques de caméras et sur des modèles d'imagerie non-centraux.



Jean **Devars**

Jean Devars est professeur à l'Université Pierre et Marie Curie (Paris 6). Il dirige l'équipe Perception et Vision du LISIF. Il travaille depuis 1985 dans le domaine de l'imagerie. C'est sur ses idées originales que l'activité 360° a débuté, il a encadré 25 thèses portant toutes sur les domaines de la vision, de l'imagerie et ce, dans différentes applications. Parallèlement, il est responsable du Mastère SDI de l'Université Pierre et Marie Curie.