

Détermination de la géométrie épipolaire d'images panoramiques cylindriques

Laurent SMADJA, Ryad BENOSMAN, Jean DEVARS, Gilles CORDURIÉ

Laboratoire des Instruments et Systèmes d'Ile de France

Université Pierre et Marie Curie

75252 Paris Cedex 05

laurent.smadja@lis.jussieu.fr, devars@ccr.jussieu.fr, Gilles.CORDURIE@cicrp.jussieu.fr

Résumé – Nous présentons ici la géométrie d'un capteur panoramique tournant, ainsi que la généralisation du concept de contrainte épipolaire sur des images de type cylindrique. Différents procédés de minimisation sont utilisés pour retrouver certains paramètres expérimentaux. Cette contrainte permet ainsi d'optimiser de l'étape de recherche de points appariés afin de reconstituer la configuration tridimensionnelle de la scène observée à partir d'images cylindriques haute définition.

Abstract – In this paper, we describe the design of a panoramic vision system. We also present a generalization of epipolar geometry for cylindrical stereo images. Knowing this constraint, minimization processes are computed to recover some experimental parameters. Furthermore, epipolar constraint optimizes the matching stage and thus, allows to achieve 3D reconstructions of outdoor static scenes from high definition images.

1. Introduction

Différents procédés peuvent être employés pour la réalisation d'images panoramiques, principalement des miroirs de forme particulière [1][2] ou des capteurs tournants. Les systèmes à miroir sont capables d'effectuer des applications temps réel, les caméras utilisées fournissent en effet un flux vidéo de 25 images par seconde, et sont généralement dévolus à la vidéo conférence ou la vision robotique[3][4][5]. Cependant, ces capteurs souffrent généralement d'un manque de résolution pénalisant pour un processus de reconstruction précise. Les capteurs omnidirectionnels basés sur des caméras tournantes [6] quant à eux, ne visent évidemment pas les applications temps réel de par la nature même de leur conception mais fournissent des images de haute résolution, ce qui permet d'atteindre la précision désirée. Cette caractéristique est très appréciable pour la reconstruction de scènes telles que des décors cinématographiques, c'est d'ailleurs l'objectif du système présenté ici. Cet article reprend les bases de [7][8] en généralisant le concept de couple stéréoscopique d'images panoramiques « cylindriques », autorisant le libre positionnement des deux caméras l'une par rapport à l'autre.

1.1 Architecture du capteur

Le capteur développé est basé sur l'utilisation d'une caméra linéaire Tri-CCD de 2048 pixels, mise en rotation sur un axe vertical 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. En concaténant les colonnes acquises au cours d'une rotation complète, on obtient une image panoramique RGB au format 2048*3600.

La mise en rotation est assurée par un ensemble constitué de :

- Une carte de commande intégrant un micro-contrôleur de la famille AVR.
- Un moteur à courant continu couplé à un encodeur optique, dont le rapport de réduction est de 1152.

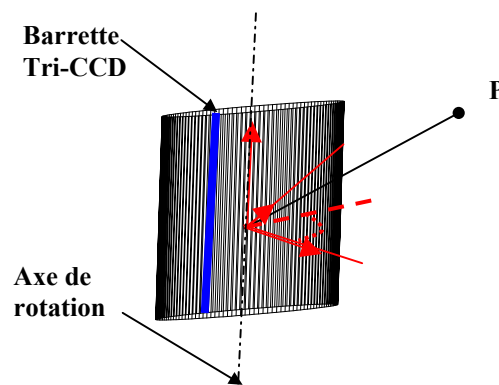


FIG 1: Géométrie du capteur

La carte reçoit des commandes de paramétrage ou de mise en rotation via une liaison série standard à 9600 bauds. Différents paramètres tels que vitesse de rotation, répartition angulaire des acquisitions et temps d'exposition sont ainsi configurables et sauvegardés dans la zone EEPROM du micro-contrôleur. Lorsque le moteur est mis en fonctionnement, cette carte délivre alors un signal modulé en largeur d'impulsion directement issu du micro-contrôleur, via un étage de commutation à base de MOSFET. L'information issue du codeur est disponible sous la forme d'un signal logique qui fournit 500 impulsions par tour du moteur, soit 576000 impulsions par tour en sortie des étages de réduction. Ce signal est utilisé pour réaliser l'asservissement du moteur

à vitesse de rotation constante, ainsi que pour assurer le déclenchement d'une acquisition vidéo.

La caméra utilisée est une *Jai CV L-103*, barrette linéaire de 2048 pixels qui délivre des images RGB à un débit maximal de 2400 images par seconde, ce qui est largement plus que nécessaire. Elle est couplée à une *Coreco Pc DIG* comme carte d'acquisition et l'objectif choisi est un *Nikon F-Or 28/80 mm*, généralement réglé sur 28 mm.

La fréquence du signal envoyé par le moteur conditionne la durée d'intégration de la barrette Tri-CCD ; il est donc choisi en fonction des conditions expérimentales d'acquisition. Par exemple, pour une scène d'extérieur, un bref temps d'exposition (entre 10 et 25 ms par colonne) est suffisant. L'acquisition des 3600 colonnes nécessaires à la formation d'une image panoramique dure alors moins d'une minute et demie. Dans le cas de scènes d'intérieur, où la luminosité est nettement plus faible, ce même temps d'exposition peut être augmenté jusqu'à 250 ms, élevant le temps d'acquisition total à près d'un quart d'heure.

La caméra est mise en rotation selon un axe passant par son centre focal, qu'il faut donc déterminer. Un procédé optique particulièrement simple dans sa mise en œuvre a été développé ; Son principe est illustré sur la figure 2. La caméra est placée entre deux plaques présentant des lignes, dont l'une est constituée pour moitié d'un miroir plan. Une acquisition est prise de manière à voir la planche munie du miroir et le reflet de la plaque opposée. En détectant sur

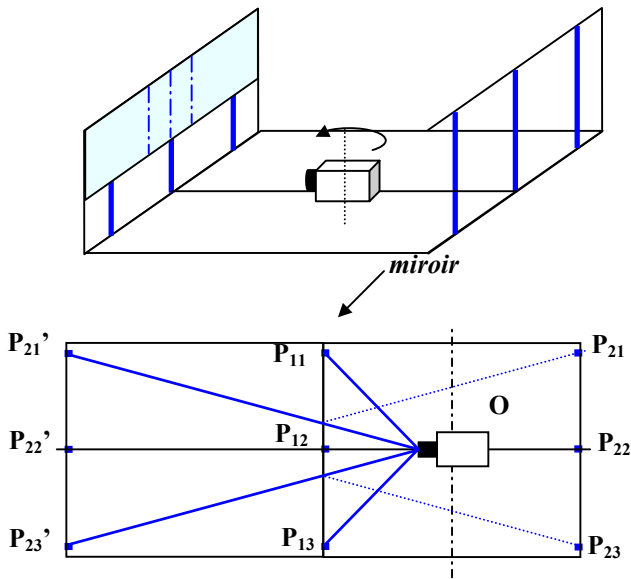


FIG. 2: Montage expérimental pour la détermination du centre optique

l'image obtenue les différentes lignes, les angles $\alpha = (\mathbf{P}_{13} \mathbf{O} \mathbf{P}_{22'})$ et $\beta = (\mathbf{P}_{21'} \mathbf{O} \mathbf{P}_{22'})$ sont recalculés. Si le centre optique de la caméra est exactement placé au milieu des deux plaques, la relation (1) doit être vérifiée.

$$\alpha = \text{Arctg}\left(\frac{1}{3} \text{tg} \beta\right) \quad (1)$$

Cette méthode est très précise et permet de placer le centre optique de la caméra sur l'axe de rotation du capteur, et ainsi de connaître le rayon du cylindre image.

La taille d'une de ces images étant très importante (environ 20 Mo en BMP), le calcul de la contrainte épipolaire s'avère fondamental pour appairer les couples d'images.

1.2 Détermination des lignes épipolaires

La quasi-totalité des systèmes de stéréoscopie existants utilisent une contrainte géométrique pour réduire le temps de calcul durant la phase d'appariement, la *contrainte épipolaire*. Cette approche permet de contraindre la recherche de points appariés sur une seule courbe de l'image au lieu de l'image entière. Si l'expression de cette contrainte est parfaitement connue pour des couples d'images perspectives classiques, peu d'auteurs l'ont étudiée dans le cas de projections inhabituelles, telles que celles fournies par les miroirs hyperboliques [9] ou des images à géométrie cylindrique [10][11].

Le modèle présenté détermine l'équation des lignes épipolaires pour un couple de deux images cylindriques prises de deux points de vue complètement arbitraires.

Soit \mathbf{P} le point physique observé, \mathbf{p}_1 et \mathbf{p}_2 ses deux images sur les deux cylindres du couple stéréoscopique. \mathbf{O}_1 et \mathbf{O}_2 sont les « foyers » respectifs de ces deux cylindres, sur lesquels sont centrés les repères orthonormés \mathbf{R}_1 et \mathbf{R}_2 . Enfin, on appelle \mathbf{T}_{21} la matrice de transformation homogène entre \mathbf{R}_1 et \mathbf{R}_2 . Dans le cas classique plan, la ligne épipolaire associée à un point de la première image est l'intersection entre le second plan image et un plan particulier, appelé *plan de vue*, constitué par \mathbf{P}, \mathbf{O}_1 et \mathbf{O}_2 . Dans notre cas, cette intersection ne sera pas une ligne mais une ellipse.

L'idée principale est ici de ne pas considérer directement \mathbf{P} mais bien ses deux images \mathbf{p}_1 et \mathbf{p}_2 . La seconde est de choisir une métrique réelle dans l'ensemble de nos calculs plutôt qu'une métrique pixellique.

Le but est ici de déterminer, dans un repère choisi, l'équation de l'intersection du plan de vue et du cylindre image. Si l'image est cylindrique alors les coordonnées de \mathbf{p}_1 dans \mathbf{R}_1 se calculent facilement :

$$p_{1O_1} \begin{cases} x_{p1} = focal * \cos(\alpha) \\ y_{p1} = focal * \sin(\alpha) \\ z_{p1} = (v_{center} - v_1) * pixelsize \end{cases} \quad (2)$$

Où :

$$\alpha = u_1 * 2\pi / Ncolumn \quad (3)$$

focal est le rayon du cylindre image, ce qui correspond à la distance focale du système de vision, v_{center} la position du centre optique dans la colonne image et *Ncolumn* le nombre de colonnes de l'image.

La transformation rigide \mathbf{T}_{12} est composée de trois rotations d'angle ζ , θ et φ autour des trois axes et de la translation entre \mathbf{O}_1 et \mathbf{O}_2 :

$$T_{12} = \begin{pmatrix} & & t_x \\ & \mathbf{R}_{12} & t_y \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (4)$$

Le même raisonnement est tenu dans le repère \mathbf{R}_2 , puis les coordonnées de \mathbf{p}_2 sont transcrites dans \mathbf{R}_1 :

$$p_{2O_1} = T_{21} * p_{2O_2} \quad (5)$$

Où \mathbf{p}_{2O_2} sont les coordonnées homogènes de \mathbf{p}_2 dans \mathbf{R}_2 . les trois points suffisants à la génération du plan de vue sont à présent connus dans le repère, l'équation de ce plan est :

$$Plane : \begin{pmatrix} x & y & z_{calcul} & 1 \\ x_{O_1} & y_{O_1} & z_{O_1} & 1 \\ x_{O_2} & y_{O_2} & z_{O_2} & 1 \\ x_{p_1} & y_{p_1} & z_{p_1} & 1 \end{pmatrix} = 0 \quad (6)$$

Le calcul de l'intersection de ce plan avec le second plan de vue donne une équation relativement compliquée, qui ne sera pas présentée ici, et dont la forme peut être modifiée jusqu'à obtenir une valeur Z_{calcul} , en fonction de la position angulaire α ($\alpha \in [0, 2\pi]$)

$$f : \alpha \rightarrow Z_{calcul}(\alpha) \quad (7)$$

La dernière étape est finalement de retraduire cette équation d'ellipse en coordonnées image, ce qui est réalisé grâce aux équations (2) et (3).

En définitive, la contrainte épipolaire cylindrique repose sur l'estimation de *focal*, v_{center} et *pixelsize* qui sont les paramètres intrinsèques relatifs à la caméra, et \mathbf{T}_{21} les paramètres extrinsèques relatifs au couple d'images. Plusieurs cas sont ainsi à étudier :

1. \mathbf{T}_{21} est inconnue, ainsi que les paramètres intrinsèques.
2. \mathbf{T}_{21} est inconnue, les paramètres intrinsèques supposés connus.

1. est instable, en effet trop de paramètres doivent être déterminés et les procédés de minimisation ne convergent pas.
2. est le cas retenu, il donne des résultats satisfaisants à condition de choisir le jeu de paramètres initial \mathbf{T}_{21} proche de la solution optimale de minimisation, ce jeu étant déterminé durant l'acquisition.

1.3 Estimation des différents paramètres

Les procédés d'optimisation se basent sur un nuage de points appariés sur les deux images cylindriques non calibrées. L'algorithme de détection et d'appariement utilisé [12] est relativement robuste car invariant en rotation, changements d'échelle et d'intensité. Les N points appariés obtenus sont utilisés pour construire une base sur laquelle les calculs de minimisation sont effectués. Nous disposons des coordonnées image de chacun des points, grâce auxquels on retrouve les coordonnées cylindre \mathbf{a}_i et $\mathbf{Z}(\mathbf{a}_i)$. Connaissant l'équation épipolaire (7), le critère à minimiser devient évident :

$$Q = \sum_{i=1}^N |z_i - z_{calcul}|^2 \quad (8)$$

Cette expression doit être dérivée par rapport à tous les paramètres que l'on détermine : ζ , θ , φ , t_x , t_y et t_z . Trois algorithmes de minimisation différents ont été utilisés. Le premier d'entre eux utilise la fonction *leastsq.m* du logiciel MATLAB, qui combine plusieurs méthodes pour converger efficacement, le second algorithme est la *descente de gradient avec pas variable*, basé sur le critère (8) et où le gradient est calculé à l'aide du logiciel MAPLE, le troisième enfin est l'algorithme de *Levenberg-Marquardt*.

2. Résultats expérimentaux

Le couple stéréo présenté sur la *figure 5* a été réalisé dans les rues d'Angoulême. La focale choisie était de 28 mm et le

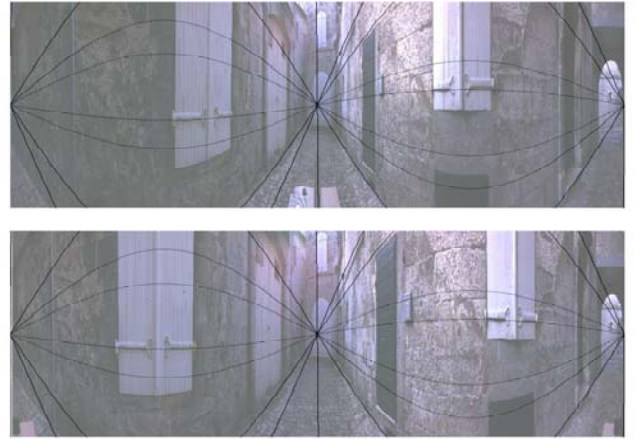


FIG. 3 : Détermination des lignes épipolaires

temps d'exposition entre 10 et 45 ms, ce temps variant selon les conditions météorologiques. On peut noter la présence de deux points particuliers dans l'image, appelés *épipoles*, qui sont les intersections du plan image et de la direction de translation ($\mathbf{O}_1\mathbf{O}_2$).

Pour une base de 45 points appariés, nous avons appliqué les trois méthodes précédemment citées ; les résultats obtenus sont résumés dans le Tableau 1

TAB. 1 résultats des minimisations

	\mathbf{T}_{12}	Leastsq	descente Gradient	Levenberg Marquardt
φ	0	0.036	- 0.015	- 0.042
θ	0	- 0.019	- 0.004	0.000
ζ	0	- 0.099	- 0.003	0.000
t_x	-1	0.153	- 0.970	- 0.910
t_y	0	0.063	- 0.324	- 0.042
t_z	0	0.141	0.039	0.003

Quelques reconstructions éparées ont été réalisées à l'aide de ce couple, l'une d'entre elles est présentée *figure 4*.

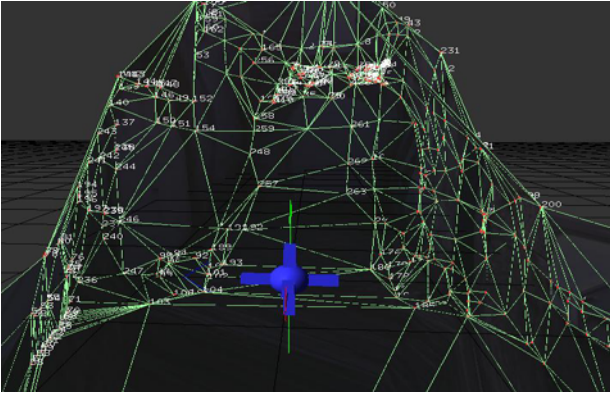


FIG 4 : Un exemple de reconstruction

3. Conclusion

Nous avons ici présenté une généralisation du concept de courbe épipolaire dans le cas d'images réalisées à l'aide de capteurs tournants. Le nouveau capteur présenté permet d'effectuer des images de ce type en un temps très raisonnable et avec une très haute définition. Plusieurs méthodes d'optimisation ont été testées et leurs différences mises en évidence.

Références

- [1] D.Southwell et al, "Panoramic stereo", I.C.P.R, pages 378-382, August 1996, Vienna, Austria.
- [2] T. Sogo, H.Ishiguro, "Real-time target localization and tracking by N-ocular stereo", pages 153-160, I.E.E.E Workshop on Omnidirectional Vision, Hilton Head Island, June 2000
- [3] E.Brassart et al., "Experimental, results got with the omnidirectional vision sensor: SYCLOP", pages 145-152, I.E.E.E Workshop on Omnidirectional Vision, Hilton Head Island, June 2000.
- [4] C.Drocourt et al., "Mobile robot localization based on a omnidirectional stereoscopic vision perception system", pages 1329-1334, I.E.E.E International Conference on Robotics and Automation, Detroit, Michigan, May 1999.
- [5] H.Ishiguro, M.Yamamoto, S.Tsuji, "Omni-directional stereo", I.E.E.E Transactions on P.A.M.I, vol. 14, No.2, pages 257-262, 1992.
- [6] Y.Pritch, M.Ben-Ezra, S.Peleg, "Automatic disparity control in stereo panoramas (Omnistere)", pages 54-61, I.E.E.E Workshop on Omnidirectional Vision, Hilton Head Island, June 12, 2000.
- [7] R.Benosman, J.Devars, "Panoramic stereovision sensor", I.C.P.R, pages 767-769, August 1998, Vienna, Austria.
- [8] R.Benosman, T.Manière, J.Devars, "Multi directional stereovision sensor, calibration and scene reconstruction", ICPR, Track A: 161-165, August 1996.
- [9] T.Svoboda, T.Pajdla, V.Hlavac, "Epipolar geometry for panoramic cameras", 5th European Conference on Computer Vision, 1998.
- [10] L.Mc Millan, "Acquiring immersive virtual environments with an uncalibrated camera", UNC Technical report 95-006
- [11] S.B Kang, R. Seliski, "3d scene recovery using multidirectional multibaseline stereo", I.J.C.V, 25(2), November 1997.
- [12] E.Bigorgne, C.Achard, J.Devars, "Local Zernike moments vector for content-based queries in image data base", Machine Vision Application, December 2000.