

Calibrage extrinsèque multi-caméras à champs non-recouvrants à l'aide d'un miroir plan

Pierre LÉBRALY, Omar AIT-AIDER, Michel DHOME, Eric ROYER

LASMEA UMR 6602, Université Blaise Pascal/CNRS
63177 Aubière Cedex, France

Pierre.LEBRALY@lasmea.univ-bpclermont.fr, Omar.AIT-AIDER@lasmea.univ-bpclermont.fr
Michel.Dhome@lasmea.univ-bpclermont.fr, Eric.Royer@lasmea.univ-bpclermont.fr

Résumé – Les systèmes multi-caméras sont utilisés dans de nombreux domaines tels que la robotique mobile ou la vidéosurveillance. Un calibrage extrinsèque précis est alors souvent nécessaire. Cette opération est généralement mise en œuvre grâce à des appariements de primitives entre différentes vues d'une même scène. Cependant, si les champs de vision des différents capteurs sont disjoints, un tel appariement est impossible à réaliser.

Cet article a pour but d'apporter une méthode de calibrage extrinsèque facile à mettre en œuvre, malgré cette contrainte. La principale contribution consiste à utiliser une scène de géométrie inconnue et de créer un champ de vision commun, entre les différents capteurs statiques, en ayant recours à un miroir plan.

Abstract – Multi-camera systems are used in many domains such as vision-based robotics or video surveillance. An accurate extrinsic calibration is usually needed. In most of cases, this task is done by matching features through different views of the same scene. However, if the sensors' field of view do not overlap, such a matching procedure isn't feasible anymore.

Despite this constraint, this article deals with an extrinsic calibration easy to implement. The main contribution consists in using a unknown geometry scene, and in using a planar mirror to create an overlap between views of the different static sensors.

1 Introduction

En fonction des applications, un système multi-caméras peut être soit statique, soit embarqué sur un véhicule. Ces travaux s'inscrivent dans cette seconde optique, la procédure de calibrage s'effectuant lorsque le véhicule est à l'arrêt. Afin de calibrer un ensemble de caméras à champs non-recouvrants, plusieurs approches ont été développées. Dans [16], l'auteur propose d'utiliser une caméra mobile supplémentaire se déplaçant à proximité des caméras à calibrer. Le calibrage est obtenu via une reconstruction 3D effectuée à un facteur d'échelle près. Une autre approche consiste à utiliser la contrainte de rigidité entre les différents capteurs du système multi-caméras en mouvement [3, 4]. L'évolution temporelle de la pose de chaque caméra est déterminée. Puis, on déduit la pose relative entre les différentes caméras, qui ne varie pas au cours du temps. Cette contrainte de rigidité peut alors servir lors de l'estimation de la pose du système multi-caméras [1, 5]. Dans [12] des connaissances *a priori* sur la vitesse du système multi-caméras et le suivi d'objet de calibrage sont combinés.

D'autre part, un miroir peut être utilisé pour obtenir des champs de vision recouvrants. L'auteur de [17] place un objet entre la caméra et des miroirs plans pour calibrer le système et reconstruire l'objet. Une autre utilisation des miroirs plans est faite dans [7] pour réaliser une tête stéréo à l'aide d'une seule caméra. Récemment ([19]), des miroirs plans ont été utilisés

pour déterminer la pose d'une mire de calibrage, par rapport à une caméra. Des travaux similaires ont été menés par [9, 10], afin de résoudre analytiquement le problème. Enfin, des travaux ont été effectués sur le calibrage extrinsèque d'un système multi-caméras, via un miroir et une mire plane connue [11].

A contrario, la principale originalité de notre approche est l'absence d'*a priori* sur la géométrie de la scène. On utilise alors des amers visuels collés sur le miroir afin d'initialiser les calculs non linéaires. Après une formalisation du problème, nous présenterons deux méthodes distinctes pour estimer les poses relatives entre différentes caméras rigidement liées. Leurs champs de vision peuvent être totalement disjoints.

2 Calibrage multi-caméras à champs non-recouvrants

2.1 Formalisation du problème

Considérons un ensemble statique d'au moins deux caméras rigidement liées, aux paramètres intrinsèques connus¹, et un miroir plan mobile. On note Π_j la $j^{\text{ème}}$ pose du miroir. La première caméra réelle C^{1r} observe différentes vues de la scène via le miroir mobile. La seconde C^{2r} observe directement cette même scène (cf. figure 1).

¹Déterminés à l'aide des méthodes développées dans [13, 14]

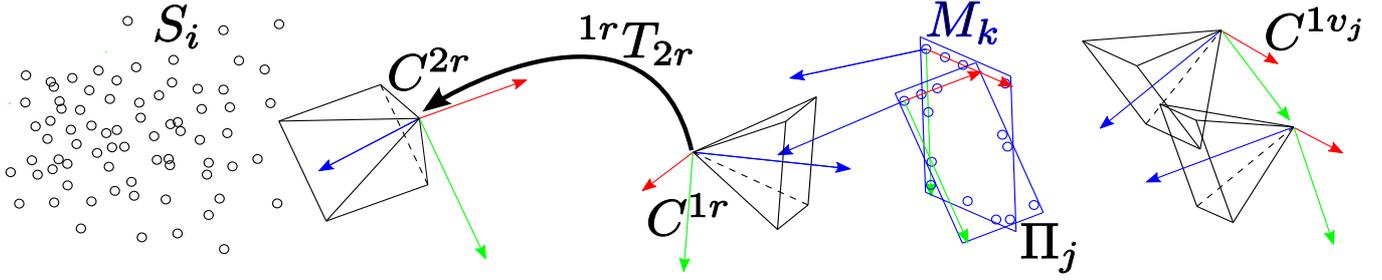


FIG. 1: Schéma du système multi-caméras et miroir

Le but est de calculer la transformation rigide entre les deux caméras, à partir des appariements 2D/2D.

On note ${}^{1r}T_{2r}$ la transformation homogène permettant de passer du repère direct de C^{1r} à celui de C^{2r} . De même ${}^{1r}T_{\Pi_j}$ désigne la transformation homogène entre le repère de C^{1r} et celui de Π_j .

On entend par *caméra virtuelle*, la caméra symétrique par rapport au plan du miroir de la caméra réelle. On note C^{1v_j} la $j^{\text{ème}}$ caméra virtuelle de C^{1r} par rapport à Π_j . On suppose que l'on a plusieurs images prises par C^{1r} et une image de la scène prise par C^{2r} .

Enfin, notons S_i les points de la scène 3D, et $s_i^{1v_j}$, s_i^{2r} leurs projections respectives dans le plan image de C^{1v_j} et de C^{2r} .

La figure 1 illustre l'ensemble de notre système. Les caméras sont représentées par leur repère, direct pour les caméras réelles, et indirect pour les caméras virtuelles. En effet, la symétrie de l'espace induite par le miroir transforme une base orthonormée directe en une base orthonormée indirecte.

Pour l'estimation précise des poses Π_j du miroir par rapport à C^{1r} , un ensemble de marqueurs M_k , approximativement connus dans le repère attaché au miroir, sont positionnés sur sa surface. On note $m_{k_j}^{1r}$ la projection des M_k dans le plan image de C^{1r} . L'utilisation de marqueurs, pastilles blanches réfléchissantes sur un fond noir, présente plusieurs intérêts :

- L'ajout de contraintes géométriques supplémentaires.
- L'estimation de la pose du miroir permet de déduire la relation entre une caméra réelle et sa caméra virtuelle. Ainsi, contrairement à [11], il n'est pas nécessaire de résoudre un système linéarisé, mais plutôt un système "exact".

L'ensemble des marqueurs M_k est estimé précisément par ajustement de faisceaux [13, 14]. Il est uniquement nécessaire de connaître précisément la distance entre deux marqueurs, afin de déterminer le facteur d'échelle² de l'ensemble du système.

Considérons une caméra réelle représentée par le modèle sté-

notopé et sa matrice de projection $K_r = \begin{pmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{pmatrix}$. En

notant $f_x = f/d_x$ et $f_y = f/d_y$, avec f la distance focale de la caméra, d_x et d_y les dimensions élémentaires de la matrice CCD et (u_0, v_0) les coordonnées pixelliques du point principal dans l'image.

Le plan image de la caméra virtuelle s'obtient par une symétrie axiale. Choisissons l'axe vertical passant par le point prin-

cipal³. En notant respectivement (u_r, v_r) et (u_v, v_v) le repère image de la caméra réelle et de la caméra virtuelle, on a alors :

$$\begin{cases} u_r = \frac{x}{d_x} + u_0 \\ v_r = \frac{y}{d_y} + v_0 \end{cases} \Leftrightarrow \begin{cases} u_v = 2u_0 - u_r = -\frac{x}{d_x} + u_0 \\ v_v = \frac{y}{d_y} + v_0 \end{cases}$$

La matrice de projection de la caméra virtuelle observant la

scène s'écrit alors $K_v = \begin{pmatrix} -f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{pmatrix}$.

2.2 Calibrage avec une mire connue

L'algorithme 1 est utilisé lorsque la géométrie des points S_i est connue.

Algorithme 1 : Calibrage extrinsèque avec une mire connue

Entrées : s_i^{2r} , $s_i^{1v_j}$, $m_{k_j}^{1r}$, M_k et S_i

Sorties : ${}^{1r}T_{2r}$, ${}^{1r}T_{\Pi_j}$, et M_k

Initialisation :

Estimation précise des marqueurs M_k .

Pour chaque image j :

 Calculer la pose du miroir Π_j dans le repère de C^{1r} .

 Calculer la pose de la caméra virtuelle C^{1v_j} par symétrie par rapport à Π_j .

 Calculer la pose de la mire S_i par rapport à C^{1v_j} .

Calculer la pose ${}^{2r}T_{S_i}$ de la caméra C^{2r} par rapport à la scène S_i . On obtient alors la pose initiale ${}^{1r}T_{2r}$ en "moyennant" les différentes estimations [8, 15, 18].

Optimisation non-linéaire :

$$\underset{({}^{1r}T_{2r}, {}^{1r}T_{\Pi_j}, {}^{2r}T_{S_i})}{\operatorname{argmin}} \left(\sum_k \|m_{k_j}^{1r} - \operatorname{proj}_{C^{1r}}(M_k)\|^2 + \sum_i \left[\|s_i^{2r} - \operatorname{proj}_{C^{2r}}(S_i)\|^2 + \sum_j \|s_i^{1v_j} - \operatorname{proj}_{C^{1v_j}}(S_i)\|^2 \right] \right)$$

Avec $\operatorname{proj}_C(M)$, la fonction projetant le point M dans le plan image de la caméra C . Dans la phase d'initialisation, la pose (R, T) des objets planaires est obtenue par estimation linéaire, puis non linéaire de l'homographie H entre l'objet et le plan

²inhérent au processus de calibration

³Cette problématique est également évoquée dans [6].

image de la caméra l'observant [20], par les équations :

$$\begin{cases} \lambda = \|K^{-1}\mathbf{h}_1\|^{-1} = \|K^{-1}\mathbf{h}_2\|^{-1} \\ T = \lambda K^{-1}\mathbf{h}_3 \end{cases} \text{ et } \begin{cases} \mathbf{r}_1 = \lambda K^{-1}\mathbf{h}_1 \\ \mathbf{r}_2 = \lambda K^{-1}\mathbf{h}_2 \\ \mathbf{r}_3 = \mathbf{r}_1 \wedge \mathbf{r}_2 \end{cases}$$

Avec $R = [\mathbf{r}_1 \ \mathbf{r}_2 \ \mathbf{r}_3]$, $H = [\mathbf{h}_1 \ \mathbf{h}_2 \ \mathbf{h}_3]$ et K la matrice de projection de la caméra considérée.

La pose des objets 3D est estimée par la méthode décrite dans [2], puis par optimisation non-linéaire⁴.

2.3 Calibrage sans mire connue

En considérant à présent que la géométrie des points S_i est inconnue, aucune mire de calibrage n'est nécessaire. Ainsi, on ne restreint plus le champ de vision à la seule mire. Tout le champ d'observation fourni par le miroir est utilisé. Le but est de rendre la procédure de calibrage plus souple et plus aisée pour l'utilisateur. L'enjeu porte alors sur la précision des points détectés dans les images. La méthode proposée est décrite par l'algorithme 2.

Algorithme 2 : Calibrage extrinsèque sans mire connue

Entrées : s_i^{2r} , s_i^{1vj} , $m_{k_j}^{1r}$ et M_k

Sorties : ${}^{1r}T_{2r}$, ${}^{1r}T_{\Pi_j}$, M_k et S_i

Initialisation :

Estimation précise des marqueurs M_k .

Pour chaque image j :

 Calculer la pose du miroir Π_j dans le repère de C^{1r} .

 Calculer la pose de la caméra virtuelle C^{1vj} par symétrie par rapport à Π_j .

 Trianguler les points de la scène S_i dans le repère de la caméra C^{1r} .

 Calculer la pose ${}^{2r}T_{S_i}$ de la caméra C^{2r} par rapport à la scène S_i . On obtient alors la pose initiale ${}^{1r}T_{2r}$.

Optimisation non-linéaire :

$$\begin{aligned} & \underset{({}^{1r}T_{2r}, {}^{1r}T_{\Pi_j}, S_i)}{\operatorname{argmin}} \left(\sum_k \|m_{k_j}^{1r} - \operatorname{proj}_{C^{1r}}(M_k)\|^2 \right. \\ & \left. + \sum_i \left[\|s_i^{2r} - \operatorname{proj}_{C^{2r}}(S_i)\|^2 + \sum_j \|s_i^{1vj} - \operatorname{proj}_{C^{1vj}}(S_i)\|^2 \right] \right) \end{aligned}$$

Dans le schéma proposé, C^{1r} et C^{2r} ne jouent pas un rôle symétrique. En effet, la quantité de données issues de C^{1r} est d'autant plus grande que le nombre de poses du miroir est élevé, alors que l'information issue de C^{1r} ne provient que d'une seule image. Les méthodes sont décrites ainsi pour des raisons de clarté. Cependant, il est très simple de les rendre symétriques en prenant en compte, dans le processus d'optimisation, de nouvelles images obtenues en plaçant le miroir devant C^{2r} .

⁴Algorithme de Levenberg Marquardt minimisant les erreurs de reprojec-tion, en représentant chaque rotation à l'aide d'un quaternion unitaire.

3 Résultats

Les méthodes présentées ont été testées avec des données synthétiques, fournissant une vérité terrain. Le protocole expérimental était le suivant : tout d'abord, on génère des points S_i de la scène. On génère également la pose de chaque caméra, ainsi que des poses aléatoires du miroir à environ 150mm de la caméra C^{1r} . Cette vérité terrain sert à calculer les points image 2D s_i^{2r} , s_i^{1vj} et $m_{k_j}^{1r}$. Puis, on leur applique un même bruit blanc additif suivant une loi gaussienne d'écart type σ . La pose relative entre les deux caméras est ensuite estimée à partir des données bruitées. Par ailleurs, on suppose que tous les points de la scène sont observables via chaque pose du miroir, et on prend 6 images de ce dernier. Les paramètres intrinsèques sont supposés parfaitement connus.

Pour ces simulations, on cherche à estimer la transformation ${}^{1r}T_{2r}$ correspondant à une translation de $(0.3 \ -0.5 \ -1)^T m$ et à une rotation déterminée par les angles de roulis, tangage, lacet de $(175^\circ, -45^\circ, 90^\circ)$.

Pour chaque niveau de bruit, 20 mesures ont été effectuées. L'erreur commise lors de l'estimation de ${}^{1r}T_{2r}$ est évaluée par la norme $\|dT\|$ de la différence entre la translation réelle et estimée, et par l'erreur angulaire dR entre la rotation réelle R^* et estimée \hat{R} . Plus précisément : $dR = \arccos\left(\frac{\operatorname{trace}(\hat{R}^T R^*) - 1}{2}\right)$

La figure 2 indique les performances de l'algorithme 1, utilisant une mire plane de 84 points, connue avec précision et placée à un mètre de C^{2r} . Cette méthode sert de base de comparaison pour l'algorithme 2 dans lequel on observe une scène 3D de 84 points, générés aléatoirement à environ 1 m de C^{2r} .

On observe que pour un même nombre de points, l'estimation de la translation est similaire pour les deux méthodes. Alors que l'estimation de la rotation est en moyenne trois fois plus précise avec la deuxième méthode proposée. En effet les points de la mire sont restreints à une zone relativement limitée de l'espace, or nos expériences montrent que plus les amers visuels de la scène sont dispersés, plus l'estimation des paramètres extrinsèques est précise. Cependant, le bruit de détection des amers naturels peut compenser la précision apportée par la dispersion. La précision obtenue dépend alors d'un compromis entre le nombre, la dispersion et le bruit de détection des amers.

4 Conclusion

Une méthode de calibrage extrinsèque, d'un ensemble de caméras à champs disjoints, a été présentée, puis testée avec des données synthétiques. Les résultats ont montré que cette méthode, plus souple à mettre en œuvre car elle n'utilise pas de mire connue, présente des performances similaires à une méthode de calibrage avec une mire connue, moyennant une dispersion suffisante des primitives de la scène. Les travaux en cours portent sur la validation expérimentale de l'approche à partir de données réelles à l'aide d'un banc de métrologie. L'un des enjeux porte sur la détermination d'amers naturels avec une précision sous-pixellique.

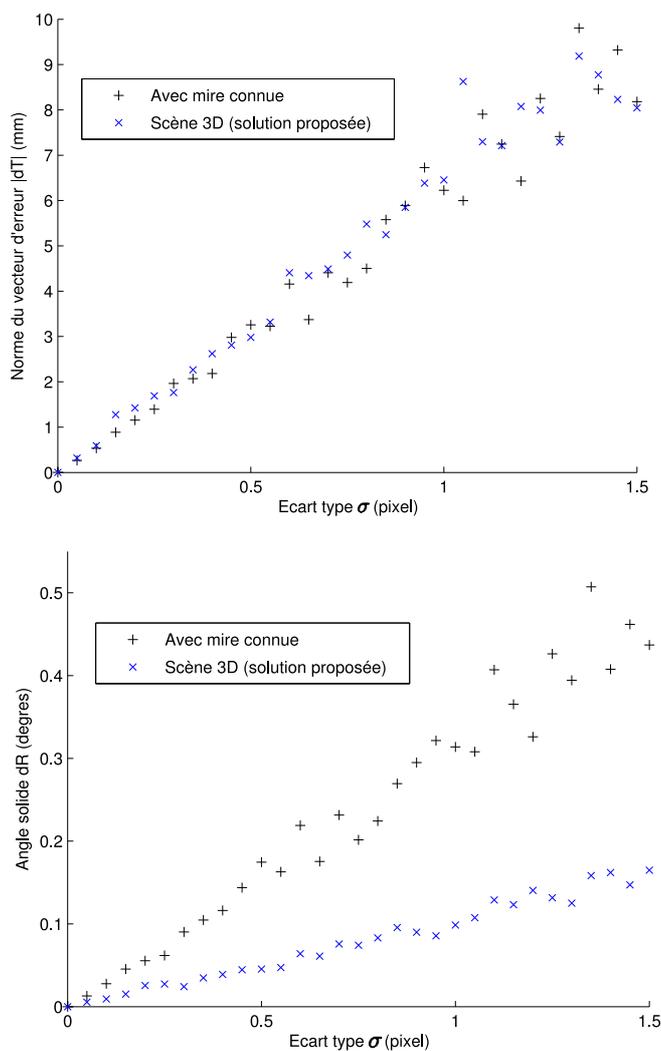


FIG. 2: Erreur moyenne sur l'estimation de pose en fonction du niveau de bruit

Références

- [1] B. Clipp, J.H. Kim, J.M. Frahm, M. Pollefeys, and R.I. Hartley. Robust 6dof motion estimation for non-overlapping, multi-camera systems. pages 1–8, 2008.
- [2] Daniel F. Dementhon and Larry S. Davis. Model-based object pose in 25 lines of code. *IJCV*, 15 :123–141, 1995.
- [3] S. Esquivel, F. Woelk, and R. Koch. Calibration of a multi-camera rig from non-overlapping views. pages 82–91, 2007.
- [4] S. Esquivel, F. Woelk, and R. Koch. Calibration of a multi-camera rig from non-overlapping views. pages 82–91, 2007.
- [5] J.M. Frahm, K. Koeser, and R. Koch. Pose estimation for multi-camera systems. pages 286–293, 2004.
- [6] A. R.J. François, G. G. Medioni, and R. Waupotitsch. Mirror symmetry \implies 2-view stereo geometry. *Image and Vision Computing*, 21(2) :137–143, February 2003.
- [7] Joshua Gluckman and Shree K. Nayar. Catadioptric stereo using planar mirrors. *Int. J. Comput. Vision*, 44(1) :65–79, 2001.
- [8] Claus Gramkow. On averaging rotations. *J. Math. Imaging Vis.*, 15(1-2) :7–16, 2001.
- [9] Joel A. Hesch, Anastasios I. Mourikis, and Stergios I. Roumeliotis. Determining the camera to robot-body transformation from planar mirror reflections. In *IROS08*, volume 1, pages 3865–3871, Nice, France, Sept 22 – Sept 26 2008.
- [10] Joel A. Hesch, Anastasios I. Mourikis, and Stergios I. Roumeliotis. Mirror-based extrinsic camera calibration. In *In The Eighth International Workshop on the Algorithmic Foundations of Robotics (WAFR'08)*, Mexico, Dec 22 – Dec 26 2008.
- [11] R.K. Kumar, A. Ilie, J.-M. Frahm, and M. Pollefeys. Simple calibration of non-overlapping cameras with a mirror. *CVPR08*.
- [12] Bernhard Lamprecht, Stefan Rass, Simone Fuchs, and Oghere Kyamakya. Extrinsic camera calibration for an on-board two-camera system without overlapping field of view.
- [13] Jean M. Lavest, Marc Viala, and Michel Dhome. Quelle précision pour une mire d'étalonnage. *Traitement du Signal*, 16(3) :241–254, 1999.
- [14] Jean-Marc Lavest, Marc Viala, and Michel Dhome. Do we really need an accurate calibration pattern to achieve a reliable camera calibration? In *ECCV '98: Proceedings of the 5th European Conference on Computer Vision-Volume I*, pages 158–174, London, UK, 1998. Springer-Verlag.
- [15] Maher Moakher. Means and averaging in the group of rotations, 2002.
- [16] Etienne Mouragnon. *Reconstruction 3D et localisation simultanée de caméras mobiles : une approche temps-réel par ajustement de faisceaux local*. PhD thesis, Ecole Doctorale Sciences Pour l'Ingénieur de Clermont-Ferrand, 2007.
- [17] Martins N. Camera calibration using reflections in planar mirrors and object reconstruction using volume carving method. *Imaging Science Journal, The*, 52 :117–130(14), 1 June 2004.
- [18] Xavier Pennec. Computing the Mean of Geometric Features Application to the Mean Rotation. Technical Report RR-3371, INRIA, 03 1998.
- [19] P.F. Sturm and T. Bonfort. How to compute the pose of an object without a direct view? pages II :21–31, 2006.
- [20] Zhengyou Zhang. A flexible new technique for camera calibration. *IEEE Transactions on PAMI*, 22(11) :1330–1334, 2000.