HAL
open science

Suivi de caméra à partir de marqueurs plans composés de cercles concentriques : paradigme et algorithmes

Lilian Calvet, Pierre Gurdjos, Vincent Charvillat, Simone Gasparini, Peter Sturm

To cite this version:

Lilian Calvet, Pierre Gurdjos, Vincent Charvillat, Simone Gasparini, Peter Sturm. Suivi de caméra à partir de marqueurs plans composés de cercles concentriques : paradigme et algorithmes. ORASIS - Congrès des jeunes chercheurs en vision par ordinateur, INRIA Grenoble Rhône-Alpes, Jun 2011, Praz-sur-Arly, France. inria-00596248

HAL Id: inria-00596248

https://inria.hal.science/inria-00596248v1

Submitted on 26 May 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Suivi de caméra à partir de marqueurs plans composés de cercles concentriques : paradigme et algorithmes

L. Calvet, P. Gurdjos, V. Charvillat [1] S. Gasparini, P. Sturm[2]

[1] IRIT - ENSEEIHT
[2] INRIA Rhône Alpes

lilian.calvet@enseeiht.fr

Résumé

Nous nous intéressons au suivi de caméra (camera tracking) à partir d'un marqueur plan composé de deux ensembles de cercles concentriques. Un ensemble de cercles concentriques (C^2Tag) a des propriétés très adaptées au suivi de caméra : un C^2Tag est très discriminant dans les images, il peut être détecté et ajusté même en cas d'occultation ; il permet aussi son identification grâce à un possible codage. De plus, sa géométrie est liée intrinsèquement aux propriétés euclidiennes de son plan de support. Nous donnons le paradigme général du suivi de deux C^2Tag coplanaires en détaillant les étapes nécessaires à sa mise en œuvre : détection, identification, ajustement, reconstruction 2D, calibrage et reconstruction 3D. Notre contribution porte sur la dernière étape consistant en un ajustement de faisceaux : nous décrivons un paramétrage minimal et un critère permettant de minimiser la distance des points extraits aux ellipses composant le marqueur. Des expériences ont été réalisées sur des données de synthèse et des données réelles et confirment la performance des algorithmes.

Mots Clef

Ajustement de faisceaux, suivi de caméra, cercles concententriques, marqueurs fiduciaux.

Abstract

In this paper we present a marker-based camera tracking with fiducial markers made up of a set of concentric circles. A set of concentric circles (C^2Tag) is well suited to cope with the camera tracking problem : a C^2Tag is very discriminatory in pictures, its detection is also possible in occluded cases and it can be identified by mean of a possible coding. In addition, its geometry is closely related to the euclidian structure of the support plane. We present the general paradigm of camera tracking based on two coplanar C^2Tag and the different steps involved : detection, identification, adjustment, 2D reconstruction, calibration and 3D reconstruction. Our contribution focuses on the last step which consists of a bundle adjustment : we describe a minimal parametrization and a criterion which minimizes the distance between measured points and the ellipses constituting the marker. Experiments on synthetic data and real data confirm the performance of the proposed method.

Keywords

Bundle adjustment, camera tracking , concentric circles, fiducial marker.

1 Introduction

1.1 Généralités

Les techniques de composition numérique (*compositing*) mélangeant images réelles d'un film et images de synthèse ont bénéficié de l'apparition de logiciels dédiés très performants [2] issus de la recherche en vision par ordinateur, en particulier de calcul de la pose de la caméra (*camera tracking*).

Le problème général. Le suivi d'une caméra consiste à calculer, à partir d'une séquence d'images, les différentes positions, orientations « internes » et « externes » de la caméra, pour *chacune* des images. L'automatisation du suivi est un point-clé du problème. Une méthode automatique se décompose en 4 étapes principales :

- Extraction des points d'intérêt dans chaque image.
- Mise en correspondance robuste des points d'intérêt entre images voisines.
- Reconstruction projective des points d'intérêt appariés.
- Reconstruction métrique (autocalibrage).

Dans le domaine du suivi basé modèle, on peut citer l'ensemble des approches basées sur des marqueurs dits « fiduciaux » (détectables et identifiables sans ambiguïté pour résumer). Parmi les plus populaires et connus les marqueurs plans d'ARToolkit [12] sont notables.

Le problème spécifique. Nous avons choisi d'utiliser des marqueurs plans formés par un ensemble de cercles concentriques (cf. le marqueur C^2Tag dans la figure 1) qui offrent beaucoup d'intérêts pour notre application si leur échelle est adaptable.

La motivation. Les marqueurs C^2Tag sont détectables, identifiables y compris en présence d'occultations et

FIGURE 1 – Image d'un marqueur C^2Tag.

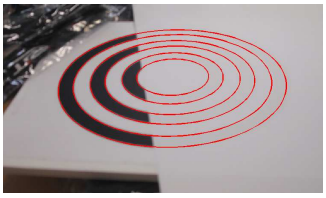


FIGURE 2 – Détection du C^2Tag en présence d'occultation. Les ellipses détectées sont mises en évidence en rouge sur l'image.

sans doute effaçables par des techniques spécifiques d'*inpainting*. Notre système de marqueur présente de bonnes performances selon les critères décrits dans [1] à savoir le taux de faux positifs, de faux négatifs, la taille minimale de l'image du marqueur et enfin son immunité au flou de bougé. De surcroît ils « codent » implicitement une partie de la structure géométrique de la scène [9]. L'évaluation du système de détection ne fait pas l'objet de cet article. La méthode de détection et d'identification est brièvement présentée dans la section 2.1.
Les différentes étapes d'un suivi de caméra à partir de marqueurs C^2Tag sont :

0. Identification de vues-clé
1. Suivi à partir des vues-clé
   (a) reconstruction 2D des C^2Tag
   (b) calibrage intrinsèque de la caméra
   (c) reconstruction 3D des C^2Tag/des caméras
   (d) ajustement de faisceaux.
2. Suivi à partir des vues intermédiaires.

L'étape 0 est une étape effectuée pour des séquences vidéo lorsque les contraintes de temps de calcul sont élevées. Elle permet d'effectuer un ajustement de faisceaux local permettant le raffinement simultané de la structure 3D de la scène et des poses des caméras le long de la séquence, ceci réduisant significativement la complexité du calcul comparé à un ajustement de faisceaux global.
L'ajustement de faisceaux est l'objet de la section §3.

1.2 Travaux antérieurs

Un estimateur au maximum de vraisemblance dans le cas d'approches basées points d'intérêt consiste à minimiser une fonction de coût donnée par le résidu de la reprojection (distance euclidienne entre les points mesurés et les points reprojetés) par rapport à tous les paramètres inconnus du modèle. Cette technique est optimale sous l'hypothèse d'un bruit gaussien centré et indépendant sur les points image donnés. Les paramètres du modèle pouvant être inclus dans cette optimisation sont, en particulier, la structure 3D, les matrices de projection et le cas échéant, les données de calibrage.
Dans [10] et [11] le terme optimal est utilisé dans le sens du maximum de vraisemblance et sous les hypothèses décrites ci-dessus. Ces méthodes ont une fonction de coût non linéaire et mettent en jeu un très grand nombre de paramètres. La phase d'optimisation itérative induit un coût de calcul et de mémoire très élevé.
La méthode d'ajustement de faisceaux [10] tire profit de la faible interaction de certains sous-groupes de paramètres en exploitant la structure creuse de la matrice jacobienne associée à la fonction de coût.
Dans [11], les estimations robustes du mouvement des caméras sont calculées en temps réel, des vues-clé sont sélectionnées afin de lancer la reconstruction 3D des points. La méthode introduit un ajustement de faisceaux local réduisant ainsi significativement les coûts de calcul.

Ces méthodes d'ajustement de faisceaux ne sont pas adaptées à notre scène puisque nous ne disposons pas de mises en correspondance des points mesurés dans les images. Ainsi, dans notre cas, le résidu de reprojection est la distance euclidienne entre les points mesurés et les ellipses, projection des cercles de l'espace 3D. En pratique, le résidu est une approximation de la distance euclidienne entre les points mesurés et les ellipses reprojetées. En effet, la distance euclidienne entre un point et une ellipse induit la résolution d'une équation quartique, celle-ci est donc habituellement remplacée par une approximation plus simple et plus efficace. Il existe de nombreuses approximations de la distance d'un point à une ellipse. Sampson [15] propose comme approximation la distance algébrique pondérée par l'inverse de la norme du gradient. Plus récemment cette approche a été revisitée par [4] qui note que Sampson fournit *la distance à l'approximation du premier ordre de la distance euclidienne* alors que ce dernier fournit *la distance à la première approximation de la distance euclidienne*. [14] propose une approximation basée sur la propriété d'orthogonalité entre les familles d'ellipses et d'hyperboles homofocales.

Rosin [13] compare les approximations [4], [15] et [14]. Il montre alors que l'approximation proposée par [4] est généralement plus satisfaisante que celle proposée dans [15] mais [14] reste la meilleure approximation.

1.3 Contributions

Cet article donne tout d'abord une description détaillée de l'intégralité de la chaîne d'un suivi automatique de caméra basé sur des marqueurs plans de type circulaire.
Le résultat principal donné dans cet article est un paramé-

trage exprimant la structure 3D d'une scène composée de deux C^2Tag sur un même plan.
Le paradigme d'un suivi de caméra à partir de marqueurs C^2Tag est présenté dans la section 2. Dans la section 3 nous décrivons un paramétrage minimal et un critère permettant de minimiser la distance des points extraits aux ellipses composant l'image du marqueur. Finalement, nous donnons nos résultats expérimentaux et nos conclusions dans les sections 4 et 5.

2 Suivi de caméra à partir de deux C^2Tag coplanaires

Nous considérons un motif plan composé de deux C^2Tag coplanaires. Notons par π le plan de support. Nous désignons par

$$\mathsf{P}^j = \mathsf{K}\left(\ \mathsf{R}^j \mid \mathbf{t}^j \ \right) \tag{1}$$

la matrice de la projection perspective liée à la vue j, où K est la matrice de calibrage, R^j et $\mathbf{t}^j$ sont respectivement l'orientation et la position de la caméra relativement au plan de support π, et par

$$\mathsf{H}^j = \mathsf{K}\left(\ \mathbf{r}_1^j \mid \mathbf{r}_2^j \mid \mathbf{t}^j \ \right) \tag{2}$$

la matrice de l'homographie du plan π vers la vue j, où $\mathbf{r}_1$ et $\mathbf{r}_2$ désignent les deux premières colonnes de R^j.

2.1 Reconstruction 2D d'un C^2Tag

Détection et ajustement des images des cercles. Le sous-problème traité ici est le suivant : *a) détecter les points correspondant aux images des cercles concentriques et ajuster les points détectés à des ellipses.* Jiang et al. [6] ont été les premiers à proposer un détecteur de cercles concentriques. Leur algorithme est itératif et sa convergence peut être prouvée grâce à une utilisation astucieuse d'invariants projectifs. Cependant leur algorithme n'est pas robuste aux occultations. Les travaux plus récents de [17] et [18] ne comportent pas cette limitation. Ces méthodes ne sont malheureusement pas adaptées à des marqueurs de petite dimension et ne tirent pas parti de l'information fournie par plusieurs cercles concentriques.

Nous avons conçu une méthode de détection tirant profit de la présence de $N \geq 4$ cercles concentriques. Les principaux intérêts d'utiliser un plus grand nombre de cercles sont :

- accélérer la localisation du motif dans l'image en évitant des pré-traitements sur les contours détectés ;
- garantir la présence sur la partie visible de l'objet d'un nombre de points mesurés suffisant afin d'estimer de façon fiable et précise les ellipses.

Les principales étapes de la méthode de détection sont :

1. extraction de contours ;
2. localisation de points contours situés sur l'ellipse interne par une procédure de vote ;

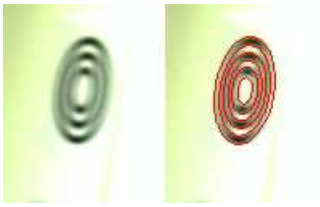


FIGURE 3 – Détection du C^2Tag en présence de flou de bougé. Les ellipses détectées sont tracées en rouge sur l'image de droite.

3. collecte des points situés sur les ellipses intermédiaires issus des résultats de l'étape 2 ;
4. procédure d'*ellipse growing* [8] pour la détermination de l'ellipse externe ;
5. ajustement des images des cercles par contraintes homologiques, similaire à l'approche [6].

Des résultats de détections sont montrés dans les figures 2 et 3. La figure 2 montre la capacité de ce type de marqueur à être détecté dans le cas de forte occultation. De plus, dans la figure 3, on peut noter que les propriétés de symétrie du C^2Tag rend la détection robuste au flou de bougé ce qui nous laisse penser qu'il existe souvent une zone non affectée par ce type de flou.

Calcul de la *structure euclidienne* du plan de support. Le sous-problème traité ici est le suivant : *pour chaque vue j, à partir des images de deux marqueurs coplanaires C^2Tag, estimer l'homographie du plan de support vers la vue j, à une similarité du plan près*. On appelle ***structure euclidienne*** du plan de support, noté π, la donnée, dans la vue j, des images des deux points cycliques [5, p52] de π, dont les vecteurs conjugués sont de la forme

$$\mathsf{H}^j \mathsf{T}_s^j \begin{pmatrix} 1 \\ \pm i \\ 0 \end{pmatrix} \ , \ i^2 = -1 \tag{3}$$

où T_s^j est la matrice d'une similarité plane arbitraire de π et H^j est la matrice de l'homographie donnée en (2). Il est ainsi clair qu'estimer la structure euclidienne de π est équivalent à estimer l'image des points cycliques.
La forme duale de cette structure est donnée par une conique tangentielle de π, appelée conique duale des points cycliques dans [5], de matrice symétrique, d'ordre 3 et de rang 2, $\mathbf{\Sigma}_\infty^j = \mathbf{h}_1^j \mathbf{h}_2^{jT} + \mathbf{h}_2^j \mathbf{h}_1^{jT}$, où $\mathbf{h}_1^j$, $\mathbf{h}_2^j$ désignent les deux premières colonnes de H^j. Le problème considéré dans notre approche consiste à *estimer $\mathbf{\Sigma}_\infty^j$ à partir des images de deux marqueurs coplanaires C^2Tag*. Chaque paire de cercles (concentriques ou supposés disjoints) fournit un ensemble de cinq équations linéaires en $\mathbf{\Sigma}_\infty^j$, dont quatre sont indépendantes [3]. Deux cercles sont choisis pour chacun des deux marqueurs, le nombre de combinaisons de deux cercles à partir de deux paires étant de quatre.
Nous estimons $\mathbf{\Sigma}_\infty^j$ en cherchant une solution $\hat{\mathbf{\Sigma}}_\infty$ au sens des moindres carrés totaux linéaires puis décomposons

$\hat{\mathbf{\Sigma}}_\infty$ [5, p55] pour obtenir la matrice d'une homographie $\hat{\mathsf{H}}^j$ qui, lorsqu'il n'y a pas de bruit de mesure, vérifie

$$\hat{\mathsf{H}}^j \sim \mathsf{H}^j \mathsf{T}_s^j \tag{4}$$

2.2 Calibrage intrinsèque de la caméra

Les points cycliques de π sont des points du cercle absolu. Nous savons [5, p211] que le calibrage intrinsèque de la caméra, paramétré par la matrice K, équivaut à déterminer l'image du cercle absolu, de matrice $\boldsymbol{\omega} = \mathsf{K}^{-T}\mathsf{K}^{-1}$, et que, pour y arriver, nous pouvons procéder « point par point », à partir des images des points cycliques, obtenues sous la forme $\hat{\mathsf{H}}^j \begin{pmatrix} 1 \\ \pm i \\ 0 \end{pmatrix}$. Les équations de base de cette approche sont les équations (8.12) données dans [5, p211]. Le problème peut être formulé de façon linéaire au sens des moindres carrés totaux sachant qu'au moins « deux vues et $1/2$ » sont nécessaires pour obtenir une solution exacte. Il existe plusieurs façons de trouver une solution pour K à partir de la solution $\hat{\boldsymbol{\omega}}$, mais la plus générale consiste à utiliser une décomposition de Cholesky de $\hat{\boldsymbol{\omega}}$ [5, p582].

2.3 Reconstruction 3D des cercles

Le problème traité est le suivant : *étant données les j homographies obtenues sous la forme (4), c.-à-d. à une similarité arbitraire près du plan de support, et la matrice de calibrage K, estimer la position et l'orientation de la caméra dans le référentiel objet.* Plus précisément, il s'agit de trouver une matrice de rotation R^j et un vecteur de translation $\mathbf{t}^j$ associés à la $j$ème caméra tels que la matrice de projection P_j s'écrive sous la forme (1).
Pour chaque vue, cela nécessite deux étapes :

1. Conformément à l'approche [16], on cherche une matrice orthogonale S^j et un vecteur $\mathbf{w}^j$ solutions de

$$\min_{\mathsf{S}^j, \mathbf{w}^j, \lambda} \| \lambda \mathsf{K}^{-1}\hat{\mathsf{H}}^j - \mathsf{S}^j \begin{pmatrix} 1 & 0 & \\ 0 & 1 & \mathbf{w}^j \\ 0 & 0 & \end{pmatrix} \|_F^2$$
$$\text{s.c. } (\mathsf{S}^j)^T \mathsf{S}^j = \mathsf{I}_3$$

   Cela permet de « corriger au mieux » la matrice $\hat{\mathsf{H}}^j$ en la remplaçant par le produit $\tilde{\mathsf{H}}^j = \mathsf{K}\mathsf{S}^j \begin{pmatrix} 1 & 0 & \\ 0 & 1 & \mathbf{w}^j \\ 0 & 0 & \end{pmatrix}$ qui garantit d'avoir la « bonne forme » de matrice. À noter que l'ambiguïté sur les deux solutions possibles est levée en utilisant l'hypothèse suivante : la caméra se déplace en restant toujours du même côté du plan.

2. On « aligne » les homographies $\tilde{\mathsf{H}}^j$ obtenues à l'étape 1, c.-à-d. on cherche T_s^j à partir de $\tilde{\mathsf{H}}^j$ telle que la décomposition

$$\tilde{\mathsf{H}}^j = \mathsf{H}^j \mathsf{T}_s^j \tag{5}$$

   soit vérifiée. Cette étape ne comporte alors aucune difficulté si on utilise les hypothèses suivantes : (a) les deux marqueurs sont identifiables (b) l'origine est le centre du premier marqueur et le centre du second marqueur a pour coordonnées cartésiennes $(d, 0, 0)$ où d est la distance (supposée connue) entre les deux centres.

Si on exprime l'inverse de T_s^j sous la forme :

$$(\mathsf{T}_s^j)^{-1} = \begin{pmatrix} s^j \mathsf{A}^j & \mathbf{b}^j \\ \mathbf{0}^T & 1 \end{pmatrix} \tag{6}$$

où s^j est un facteur d'échelle et A^j une matrice 2×2 de rotation, la $j$ème caméra est alors décrite par la matrice de projection suivante :

$$\mathsf{P}^j \sim \mathsf{K}\mathsf{S}^j \begin{pmatrix} \mathsf{I} \mid \mathbf{w}^j \end{pmatrix} \begin{pmatrix} s^j \mathsf{A}^j & 0 & \mathbf{b}^j \\ \mathbf{0}^T & s^j & 0 \\ \mathbf{0}^T & 0 & 1 \end{pmatrix} \tag{7}$$

avec, pour faire le lien avec (1),

$$\mathsf{R}^j = \mathsf{S}^j \begin{pmatrix} \mathsf{A}^j & 0 \\ \mathbf{0}^T & 1 \end{pmatrix} \quad \text{et} \quad \mathbf{t}^j = \frac{1}{s_j} \mathsf{S}^j \left(\begin{pmatrix} \mathbf{b}^j \\ 0 \end{pmatrix} + \mathbf{w}^j \right) \tag{8}$$

3 Ajustement de faisceaux

3.1 Formulation du problème

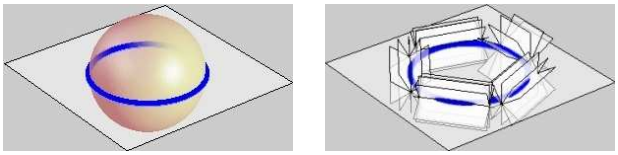
L'idée de l'ajustement de faisceaux est d'optimiser simultanément la trajectoire de la caméra et la structure de la scène. Il s'agit d'un problème aux moindres carrés non-linéaire. L'ensemble des paramètres ν est décrit dans la section 3.4. Dans notre cas, l'expression de la fonction de coût est la suivante :

$$\mathcal{C}(\nu) = \sum_j \sum_{k_1, k_2} d(x_{k_1}^j, \mathsf{E}_{j_1}) + d(x_{k_2}^j, \mathsf{E}_{j_2}) \tag{9}$$

avec E_{j_p} la projection du cercle 3D associé au marqueur p, $x_{k_p}^j$ un point mesuré sur l'ellipse externe du marqueur p, j l'indice de la vue et où $d(.,.)$ est une approximation de la distance euclidienne d'un point à une ellipse décrite dans la section 3.5.

3.2 Paramétrage du modèle

Dans cette section, nous proposons un paramétrage 3D de deux cercles coplanaires de même rayon r fixé et dont la distance inter-centres d est aussi fixée.



(a) (b)

FIGURE 4 – (a) Dans l'espace projectif 3D, un cercle correspond à un lieu de points correspondant à l'intersection d'une sphère et d'un plan ; (b) dans l'espace dual, un cercle est une quadrique dégénérée, c'est à dire une envelope de plans 3D.

Un cercle est une quadrique dégénérée. En premier lieu, concernant la modélisation suivante d'un cercle de l'espace 3D, nous utilisons une des « bizarreries » de l'espace projectif dual 3D : une conique est une quadrique tangentielle [1] dégénérée, c.-à-d. une quadrique dont la matrice d'ordre 4 est de rang 3. Cette « dualité » est illustrée dans la figure 4.
L'intérêt d'un tel modèle réside dans le fait que l'équation de projection d'une conique s'écrit matriciellement, dans sa forme tangentielle

$$\mathsf{E}_c^* \sim \mathsf{P}\mathsf{Q}_c^*\mathsf{P}^T$$

où P est la matrice de projection, Q_c^* la matrice de la quadrique tangentielle représentant le cercle et E_c^* la matrice de l'ellipse dans l'espace projectif dual, projeté de Q_c^*, sachant que la matrice de l'image du cercle se déduit grâce à la relation

$$(\mathsf{E}_c^*)^{-1} \sim \mathsf{E}_c = \begin{pmatrix} (\mathbf{e}_2 \wedge \mathbf{e}_3)^T \\ (\mathbf{e}_3 \wedge \mathbf{e}_1)^T \\ (\mathbf{e}_1 \wedge \mathbf{e}_2)^T \end{pmatrix} \tag{10}$$

où $\mathbf{e}_k$ est la colonne k de E_c^*.
On déduit facilement par dualité de [5, p202] qu'un cercle de rayon r et de centre C, obtenu par intersection d'une sphère S de rayon r et de centre C et d'un plan π passant C, est une quadrique tangentielle dégénérée, dont la matrice (d'ordre 4 et de rang 3) admet la décomposition suivante :

$$\mathsf{Q}_c^* = (\boldsymbol{\pi}^T\mathsf{S}^*\boldsymbol{\pi})\mathsf{S}^* - (\mathsf{S}^*\boldsymbol{\pi})(\mathsf{S}^*\boldsymbol{\pi})^T \tag{11}$$

où $\boldsymbol{\pi}$ est le vecteur représentant le plan π dans l'espace projectif dual et

$$\mathsf{S}^* = \mathbf{C}\mathbf{C}^T - r^2\,\mathrm{diag}(1,1,1,0) \tag{12}$$

est la matrice de S, avec $\mathbf{C}$ désignant le vecteur de coordonnées cartésiennes augmentées de C. On montre aisément qu'un tel modèle de cercles 3D a un nombre minimal de ddl, à savoir six.

3.3 Nombre minimum de paramètres

Deux cercles coplanaires, de même rayon r fixé et de distance inter-centres d aussi fixée, ont six ddl dans l'espace 3D. Nous justifions cette affirmation dans ce qui suit en associant six paramètres aux six ddl.

Centres des sphères. Soient deux sphères S_1 et S_2 centrées en C_1 et C_2 respectivement. Le centre C_1 (*3 ddl*) est paramétré par le vecteur homogène $\mathbf{C}_1$ de ses trois coordonnées cartésiennes augmentées et le centre C_2 (*2 ddl*) est paramétré par le point à l'infini C_∞ tel que

$$\mathbf{C}_2 = \mathbf{C}_1 + d\begin{pmatrix}\mathbf{u}\\0\end{pmatrix} \tag{13}$$

où $\mathbf{u} \in \mathbb{R}^3$ est le vecteur unitaire représentant la direction associée à C_∞, c.-à-d. la direction de la droite portée par C_1 et C_2 ; d est la distance (supposée connue) entre les deux centres. Les deux ddl de $\mathbf{u}$ correspondent aux deux angles θ et ϕ associés à ses coordonnées sphériques :

$$\mathbf{u} = \begin{pmatrix} cos\theta cos\phi \\ sin\theta cos\phi \\ sin\phi \end{pmatrix} \tag{14}$$

À noter que par construction, $\|\mathbf{u}\| = 1$.

Plan de support. Soit l_C la droite passant par les centres C_1 et C_2 des deux sphères et soit L_C la matrice de Plücker associée, obtenue à partir des vecteurs $\mathbf{C}_1$ et $\mathbf{C}_2$ de C_1 et C_2, via la décomposition [5, p70] $\mathsf{L}_C = \mathbf{C}_1\mathbf{C}_2^T - \mathbf{C}_2\mathbf{C}_1^T$. Le faisceau de plans ayant pour axe cette droite contient π, plan de support du cercle. Il contient aussi

- le plan π_1 défini par l_C et le centre O de la caméra, dont le vecteur est $\boldsymbol{\pi}_1 = \mathsf{L}_C^*\mathbf{O}$ avec L_C^* la représentation duale de la matrice de Plücker associée à L_C,
- le plan π_2 défini par l_C et par un autre point P non situé sur π_1, dont le vecteur est $\boldsymbol{\pi}_2 = \mathsf{L}_C^*\mathbf{P}$.

Un faisceau de plans est une famille linéaire de plans à 1 ddl. Le plan du support π (*1 ddl*) est ainsi paramétré par $\alpha \in \mathbb{R}$ tel que

$$\boldsymbol{\pi} = \alpha\boldsymbol{\pi}_1 + \boldsymbol{\pi}_2 \tag{15}$$

À noter que nous faisons implicitement l'hypothèse que le plan de support ne passe pas par le centre de la caméra car, dans (15), il n'existe pas de valeurs de α pour laquelle $\boldsymbol{\pi} = \alpha\boldsymbol{\pi}_1$.

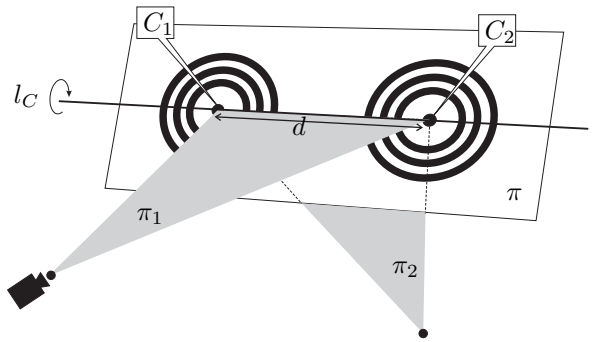


FIGURE 5 – Structure de la scène : l_C la droite passant par les centres C_1 et C_2. Le plan π_1 est défini par l_C et le centre O de la caméra. Le plan π_2 est défini par l_C et par un point non situé sur π_1.

En faisant la somme de tous les ddl mentionnés entre parenthèses, nous obtenons bien un total égal à 6.

3.4 Ensemble des paramètres

Le paramétrage est réuni dans le vecteur

$$\nu = \left[\ \nu_{in} \quad \nu_\pi^1 \quad \ldots \quad \nu_\pi^j \quad \ldots \quad \nu_\pi^N\ \right]^T \tag{16}$$

1. On appelle *quadrique tangentielle* toute quadrique de P(E*).

avec ν_{in} le vecteur des paramètres intrinsèques de la caméra et

$$\nu_{\pi}^{j} = \begin{bmatrix} \mathbf{C}_1^j \\ \phi^j \\ \theta^j \\ \alpha^j \end{bmatrix}$$

le vecteur des paramètres de la structure du plan associé à la $j$ème vue. Le vecteur de paramètres est initialisé grâce à la méthode proposée dans la section 2.

3.5 Distance point/ellipse

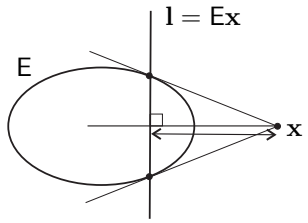


FIGURE 6 – La relation pôle-polaire : la ligne $\mathbf{l} = \mathsf{E}\mathbf{x}$ est la polaire du point $\mathbf{x}$ respectivement à la conique E [5, p58].

La fonction de coût (9) dépend du type d'approximation de la distance euclidienne d'un point à une ellipse choisi. Nous choisissons, pour sa simplicité analytique, une variante de l'approximation proposée par [7]. Nous l'interprétons géométriquement comme la distance du point de vecteur $\mathbf{x}$ à la droite polaire de ce point par rapport à l'ellipse de matrice E (cf. [5, p58] pour les relations de polarité et figure 6) :

$$d(\mathbf{x}, \mathsf{E}) = \frac{\mathbf{x}^T \mathsf{E}\mathbf{x}}{\| \, diag(1,1,0)\mathsf{E}\mathbf{x} \, \|} \tag{17}$$

4 Résultats expérimentaux

4.1 Données synthétiques

Nous présentons dans un premier temps les résultats obtenus dans la phase d'initialisation linéaire pour le calcul des paramètres intrinsèques et des positions/orientations des caméras. Dans un deuxième temps nous exposons les résultats de l'algorithme d'optimisation non-linéaire pour l'ajustement de faisceaux.

La phase d'optimisation pour l'ajustement de faisceaux est conduite via l'algorithme de Levenberg-Marquardt.

Les résultats sont exprimés via la racine carrée de l'erreur quadratique moyenne

$$REQM = \sqrt{\frac{1}{n} \sum_i (\hat{x}_i^{(k)} - x_i)^2} \tag{18}$$

avec n le nombre de paramètres estimés, $\hat{x}_i^{(k)}$ le $i$ème paramètre calculé à l'itération k et x_i le paramètre exact associé. Les figures 8, 9 et 10 présentent l'évolution de la REQM au cours des itérations des optimisations. Nous avons généré 100 jeux de données de synthèse, chacun composé d'un ensemble de vues issues de 20 caméras.

Evaluation de l'initialisation. Soit $\mathbf{x}^{(0)}$ le vecteur de paramètres initial calculé comme présenté dans la section 2. Les données initiales sont les points 2D situés sur chacune des ellipses auxquels nous avons ajouté un bruit blanc gaussien dont l'écart type est $\sigma_1 = 0.5$, $\sigma_2 = 1$ et $\sigma_3 = 2$ pixels respectivement pour la première, deuxième et troisième expérience. La figure 7 présente la REQM du vecteur $\mathbf{x}^{(0)}$ suivant le nombre de points mesurés sur chacune des ellipses.

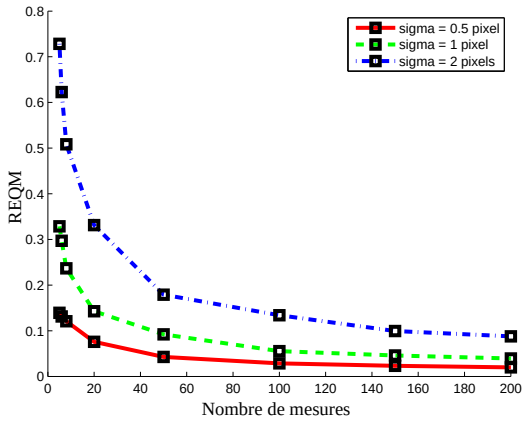


FIGURE 7 – Erreurs moyennes du vecteur $\mathbf{x}^{(0)}$ suivant le nombre de points considérés sur chacune des ellipses. L'erreur est une moyenne estimée sur 100 simulations. L'expérience est renouvelée en faisant varier l'écart type du bruit gaussien appliqué sur les mesures.

Première expérience. Dans un premier temps, l'ajustement de faisceaux a été effectué à partir du vecteur initial $\mathbf{x}^{(0)}$ issu de la procédure d'initialisation effectuée avec le même nombre de points que ceux considérés pour l'ajustement de faisceaux. L'expérience est renouvelée en faisant varier l'écart type du bruit blanc gaussien appliqué aux mesures. Les résultats sont présentés dans la figure 8. La figure 9 présente les performances de l'algorithme lorsque le nombre de points mesurés varie avec un écart type du bruit appliqué sur les mesures de 1 pixel.
Notons que le gain en précision obtenu via l'algorithme de minimisation est principalement atteint au bout de 4 itérations.

Seconde expérience : initialisation peu précise. Une deuxième expérience a été réalisée en effectuant l'ajustement de faisceaux sur un vecteur initial calculé à partir de 5 points approximativement équidistants sur chacunes des ellipses et l'ajustement de faisceaux étant effectué sur un ensemble de 200 points mesurés par ellipses. On étudie le comportement de l'ajustement de faisceaux lors d'une ini-

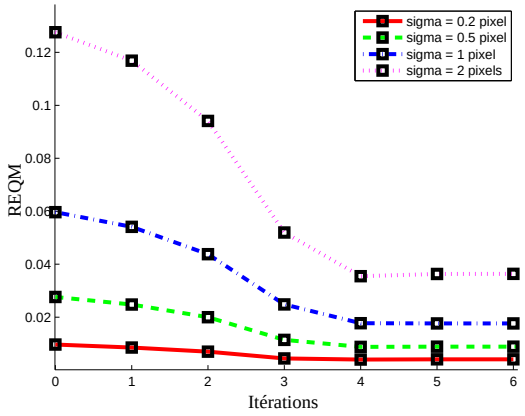


FIGURE 8 – Convergences moyennes de l'algorithme de minimisation non-linéaire sur l'ajustement de faisceaux effectué sur 20 vues. L'erreur est une moyenne estimée sur 100 simulations. L'expérience est réalisée sur 20 vues en considérant 100 points par ellipse. Celle-ci est renouvelée en faisant varier l'écart type du bruit gaussien appliqué sur les mesures.

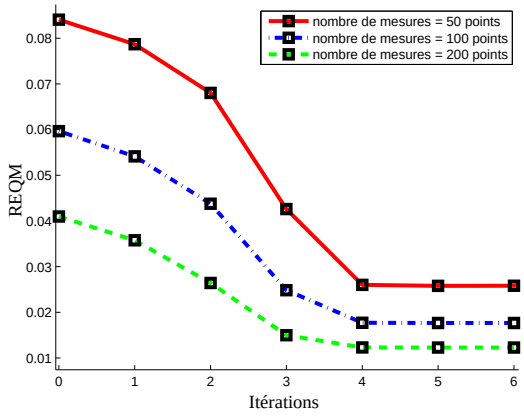


FIGURE 9 – Convergences moyennes de l'algorithme de minimisation non-linéaire sur l'ajustement de faisceaux. L'erreur est une moyenne estimée sur 100 simulations. L'expérience est réalisée sur 20 vues avec un écart type du bruit gaussien appliqué sur les mesures $\sigma = 1$ pixel. Celle-ci est renouvelée en faisant varier le nombre de points mesurés.

tialisation peu précise.

4.2 Images réelles

Sélection des vues-clé. La sélection des vues-clé est effectuée selon le critère suivant :

$$\rho = |1 - (\mathbf{l}_\infty^k)^T \mathbf{l}_\infty^j| \tag{19}$$

où $\mathbf{l}_\infty^k$ représente la droite à l'infini du plan de la dernière vue-clé sélectionnée et $\mathbf{l}_\infty^j$ représente la droite à l'infini du plan de l'image courante, toutes deux normalisées pour avoir une norme unitaire. Dès que ρ est supérieur à $\epsilon = 10^{-8}$ (déterminé de façon empirique) alors la vue $j-1$ est sélectionnée comme nouvelle vue-clé .
La figure 11 montre la sélection des vues-clé lors d'une séquence de 200 images et l'évolution du critère défini ci-dessus. La figure 12 montre pour cette même expérience les vues-clé sélectionnées, présentées par ordre chronologique.

L'ensemble de la procédure de suivi de caméra a été illustré à partir d'une vidéo (accessible à l'URL : `http://gurdjos.perso.enseeiht.fr/demos/orasis2011`). Dans cette expérience la sélection des vues-clé est effectuée. Les matrices caméra pour les vues intermédiaires sont raffinées grâce à la méthode proposée dans [3].

5 Conclusion et perspectives

Nous avons présenté dans cet article le paradigme d'un suivi de caméra à base de marqueurs constitués de cercles concentriques au sein duquel le choix pour ce type de marqueur a été justifié. Notre contribution concerne principalement la dernière étape consistant en un ajustement de faisceaux : nous avons décrit un paramétrage minimal et un critère permettant de minimiser la distance des points extraits aux ellipses composant le marqueur. Des résultats expérimentaux, sur données synthétiques et images réelles montrent les performances de la méthode.
Nous travaillons actuellement sur l'évaluation de notre algorithme lorsque l'erreur de reprojection est mesurée via d'autres approximations de la distance d'un point à une ellipse (notamment celle présentée dans [13]). De plus nous envisageons de généraliser la méthode proposée pour $N \geq 2$ C^{2}Tag non coplanaires.

Remerciements

Ce travail a bénéficié du soutien financier de l'Agence Nationale de la Recherche (ANR) dans le cadre de l'appel d'offres « Contenus et Interactions » pour le projet « *Real time On-set Matchmoving* » (ANR-08-CORD-020), dont les partenaires sont Duran, l'INRIA Rhône Alpes et l'IRIT.

Références

[1] M. Fiala. Artag, a fiducial marker system using digital techniques. In *CVPR*, pages 590–596, 2005.

[2] A. W. Fitzgibbon and A. Zisserman. Automatic camera tracking. In M. Shah and R. Kumar, editors, *Video Registration*, chapter 2, pages 18–35. Kluwer, 2003.

[3] P. Gurdjos, P. Sturm, and Y. Wu. Euclidean Structure from N>=2 Parallel Circles : Theory and Algorithms. In *ECCV*, volume 1, pages 238–252, 2006.

[4] M. Harker and P. O'Leary. First order geometric distance (the myth of sampsonus). In *BMVC*, pages 87–96, 2006.

[5] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, second edition, 2004.

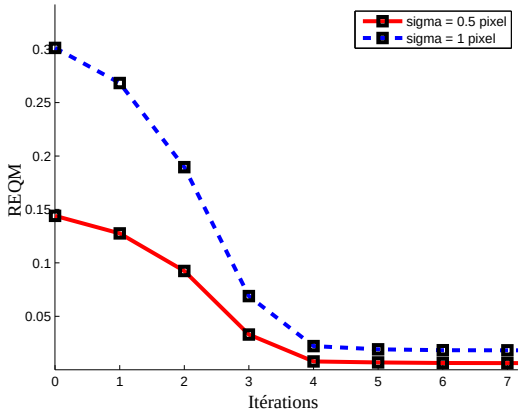


FIGURE 10 – Convergences moyennes de l'algorithme de minimisation non-linéaire sur des ajustements lorsque l'initialisation des paramètres est peu précise. L'erreur est une moyenne estimée sur 100 simulations. L'expérience est réalisée sur 20 vues et l'on considère 200 points par ellipse pour l'ajustement de faisceaux. Celle-ci est renouvelée en faisant varier l'écart type du bruit gaussien appliqué sur les mesures.

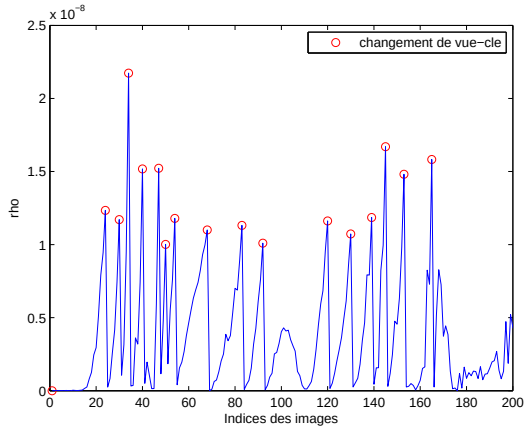


FIGURE 11 – Évolution de ρ au cours de la séquence vidéo.

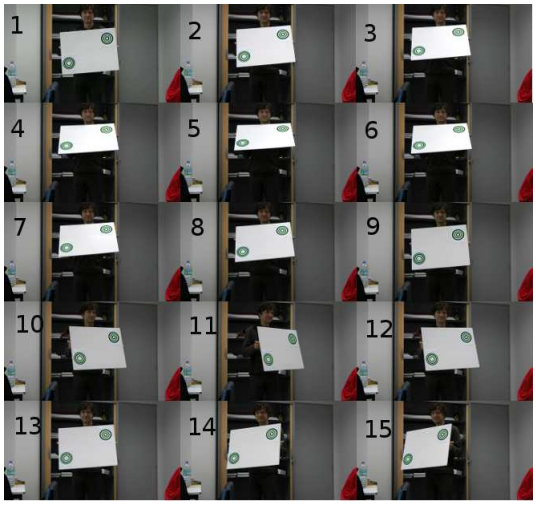


FIGURE 12 – Vues-clé sélectionnées.

[6] G. Jiang and L. Quan. Detection of concentric circles for camera calibration. In *ICCV*, pages 333–340, 2005.

[7] K. Kanatani. *Statistical Optimization for Geometric Computation : Theory and Practice*. Elsevier, Amsterdam, 1996.

[8] K. Kanatani and N. Ohta. Automatic detection of circular objects by ellipse growing. *IJIG*, 4(1) :35–50, 2004.

[9] J. Kim, P. Gurdjos, and I. Kweon. Geometric and algebraic constraints of projected concentric circles and their applications to camera calibration. *PAMI*, 27(4) :637–642, 2005.

[10] M. I. A. Lourakis and A. A. Argyros. Sba : A software package for generic sparse bundle adjustment. *ACM Trans. Math. Softw.*, 36(1), 2009.

[11] E. Mouragnon, M. Lhuillier, M. Dhome, F. Dekeyser, and P. Sayd. Generic and real-time structure from motion using local bundle adjustment. *IVC*, 27(8) :1178–1193, 2009.

[12] I. Poupyrev, H. Kato, and M. Billinghurst. *ARToolkit User Manual, Version 2.33*, 2000.

[13] P. L. Rosin. Ellipse fitting using orthogonal hyperbolae and stirling's oval. *GMIP*, 60(3) :209–213, 1998.

[14] P. L. Rosin. Evaluating harker and o'leary's distance approximation for ellipse fitting. *PRL*, 28(13) :1804–1807, 2007.

[15] P. D. Sampson. Fitting conic sections to very scattered data. an iterative refinement to the bookstein algorithm. *CVGIP*, (18) :97–108, 1982.

[16] P. Sturm. Algorithms for plane-based pose estimation. In *CVPR*, pages 1010–1017, June 2000.

[17] X. Ying and H. Zha. An efficient method for the detection of projected concentric circles. In *ICIP*, pages 560–563, 2007.

[18] X. Ying and H. Zha. Efficient detection of projected concentric circles using four intersection points on a secant line. In *ICPR*, pages 1–4, 2008.