Méthodes et fondements géométriques pour l'autocalibrage plan

Peter Sturm
INRIA Rhône-Alpes, 655 Avenue de l'Europe
38330 Montbonnot, France
www.inrialpes.fr/movi/people/Sturm

Pierre Gurdjos
IRIT-TCI, 118 Route de Narbonne
31062 Toulouse Cedex 4, France
www.irit.fr/~Pierre.Gurdjos

Résumé

Le problème traité ici est celui de l'autocalibrage d'une caméra à focale variable, à partir des vues d'une figure plane « inconnue ». Spécifiquement, les données sont limitées aux homographies inter-vues, entre une vue-clef et une vue quelconque. Il s'agit de calculer, à la fois, les paramètres intrinsèques de la caméra et ceux liés à la la structure euclidienne du plan observé. La formulation usuelle de ce problème est non linéaire et présente un inconvénient connu : l'initialisation des distances focales, requise pour une résolution itérative. Notre contribution se situe à plusieurs niveaux. (i) Nous proposons une nouvelle formulation non linéaire qui est indépendante de la distance focale. (ii) Nous proposons des solutions exactes, dont une peut être obtenue par estimation linéaire, dans le cas où le plan de la figure est un plan fronto-parallèle (relativement à la vue-clef) ; cette hypothèse peut être « relâchée » et la solution affinée par une résolution itérative. (iii) Nous établissons un lien entre certaines propriétés projectives et les principes sous-jacents à notre méthode. Nous montrons que notre approche est duale à celle usuelle, fondée sur les propriétés des images des points cycliques. En effet, la contrainte que nous proposons exprime l'alignement de trois points, dont un est le pôle de la droite de l'infini par rapport à l'image de la conique duale aux points cycliques. Les expérimentations réalisées valident la performance de nos algorithmes, en termes de précision et de complexité.

Abstract

We consider the problem of camera self-calibration, from images of a planar object with unknown Euclidean structure. The general case of possibly varying focal length is addressed. This problem is non-linear in general. One of our contributions is a non-linear approach, that makes abstraction of the (possibly varying) focal length, resulting in a computationally efficient algorithm. In addition, it does not require a good initial estimate of the focal length, unlike previous approaches. As for the initialization of other parameters, we propose a practical approach, that simply requires to take one image in roughly fronto-parallel position. Closed-form solutions for various configurations of unknown intrinsic parameters are provided. Our methods are evaluated and compared to previous approaches, using simulated and real images. Besides our practical contributions, we also provide a detailed geometrical interpretation of the principles underlying our approach.

1 Introduction

Le calibrage consiste à « retrouver » les propriétés métriques d'une caméra, à partir d'un certain nombre de vues d'une scène observée. Ces propriétés s'identifient, en partie, à un ensemble de paramètres dits intrinsèques (cf. §2 - modèle de caméra). Dans ce qui suit, par calibrage, nous sous-entendons calcul de ces paramètres intrinsèques. Les connaissances et hypothèses nécessaires à la mise en œuvre d'une méthode de calibrage concernent soit la caméra, soit la scène observée. Vis-à-vis de ce dernier point, il est important de distinguer deux classes de méthodes qui sont fondées sur des postulats diamétralement opposés : l'espace de la scène est muni d'une structure euclidienne (SE), ou ne l'est pas. Le premier postulat implique qu'un repère euclidien arbitraire peut être attaché à la scène ; on dit alors que la SE de la scène est connue. Le terme calibrage (resp. autocalibrage) s'applique dans le cas d'une scène de SE connue (resp. inconnue), qu'elle concerne un espace de dimension 1 (droite) [12], de dimension 2 (plan) [2][8][11] (resp. [5][10]) ou de dimension 3 [4].

Dans ce travail, nous supposons avoir réalisé l'acquisition de n vues d'une figure plane χ, située dans un certain plan π. Chaque vue correspond à l'image de χ par projection centrale, pour un certain paramétrage intrinsèque de la caméra dépendant d'une distance focale f spécifique (c.-à-d. variable lors des acquisitions) et de trois autres quantités supposées constantes (cf. §2). Les données dont nous disposons sont les matrices de homographies inter-vues, entre une vue-clef et une vue quelconque. Le problème à résoudre est celui de l'autocalibrage de la caméra ; il s'agit de calculer simultanément les $3 + n$ paramètres intrinsèques de la caméra (c.-à-d. calibrer) et les 4 paramètres liés à la SE du plan observé. Ces calculs sont ici indissociables ; de la relation entre ces paramètres sont issues les $2n$ contraintes usuelles de calibrage (cf. Éq. (6)). En effet, les fondements géométriques de la formulation usuelle du problème [4, §18.7][10] utilisent les propriétés suivantes. D'une part, nous savons que le paramétrage intrinsèque de la caméra est intimement lié à la représentation de l'image de la conique absolue, située sur le plan de l'infini [4, §7.5]. D'autre part, la structure géométrique de π (projective, affine, euclidienne) se déduit de la repré-

sentation de ses points cycliques (cf. §2). Les points cycliques (PC) sont deux points imaginaires conjugués [4, §1.7.3], situés à l'intersection de π et de la conique absolue (CA) ; ainsi, dans chacune des vues, les images des PC (IPC) sont aussi situées sur l'image de la CA (ICA). Ces relations entre ICA et IPC peuvent être mises sous la forme d'un système de $2n$ équations (6) ; la résolution de celui-ci (au sens des moindres carrés) fournit une solution aux problèmes du calibrage/autocalibrage plan.

L'autocalibrage plan est un problème d'estimation non linéaire de paramètres (cf. §4.2) et présente (au moins) un inconvénient connu : l'initialisation de ceux-ci et, en particulier, des distances focales, requise avant toute résolution itérative. Un autre inconvénient doit être mentionné dans le cas où le nombre d'inconnues est maximum (c.-à-d. pour n distances focales différentes) : si n croît, alors la sensibilité de l'algorithme à l'initialisation des inconnues ainsi que le temps de calcul (la complexité algorithmique est de $\mathcal{O}(r^3)$ par itération, où r est le nombre d'inconnues) augmentent de façon significative.

1.1 Contribution

Notre contribution se situe à plusieurs niveaux.

(*i*) Nous proposons une nouvelle formulation non linéaire (cf. §4.2) qui est indépendante de la distance focale. Il s'ensuit que l'algorithme proposé (cf. TAB. 1) a une très faible complexité ($\mathcal{O}(n)$). De plus, nous utilisons le fait que, si la distance focale est la seule inconnue, la formulation proposée devient indépendante des paramètres intrinsèques. Cette formulation peut être alors utilisée pour le calcul des (seuls) 4 paramètres liés à la SE du plan π, c.-à-d. pour l'(auto-)*rectification métrique* de l'image de la figure plane, ce qui généralise l'approche décrite dans [4, §1.7.5].

(*ii*) Nous établissons un lien entre certaines propriétés projectives et les principes sous-jacents à notre méthode (cf. §3.2). Notre approche est duale à celle usuelle, fondée sur les propriétés des images des points cycliques (IPC) (cf. §2). En effet, la contrainte que nous proposons exprime l'alignement de trois points, dont un est le pôle de la droite de l'infini par rapport à l'image de la conique duale aux points cycliques (ICDPC) (cf. §2 et FIG. 2).

Jusqu'à présent, nous n'avons défini que deux cas : la structure euclidienne du plan observé est connue ou non ; dans le dernier cas, nous savons que les images d'au moins quatre points (en configuration générale) permettent d'attacher un repère projectif au plan observé π ; nous dirons que la structure projective (SP) de π est connue.

(*iii*) Nous proposons des solutions exactes, dont une est obtenue par estimation linéaire, dans le cas où le plan de la figure est un plan fronto-parallèle (relativement à la vue-clef) (cf. §4.3) ; cette hypothèse peut être « relâchée » et la solution affinée par une résolution itérative. En effet, la représentation que nous proposons en §2 pour les IPC et l'ICDPC permet une stratification géométrique des problèmes à résoudre. Le calibrage plan nécessite la SE du plan observé π tandis que l'autocalibrage plan nécessite la SP de celui-ci (via les homographies inter-vues). L'hypothèse de « fronto-parallélisme » correspond à un cas intermédiaire : celui où la structure affine de π est canonique (c.-à-d. la ligne de fuite est la droite de fuite). Nous montons que ceci réduit le nombre d'inconnues de 4 à 1. Ce problème reste un problème d'estimation non linéaire, mais nous expliquons comment des solutions exactes peuvent être obtenues, en rajoutant des contraintes sur la caméra (à savoir les paramètres intrinsèques). Cette approche permet d'obtenir des valeurs initiales pour les paramètres (en évitant le problème de l'initialisation mentionné plus haut) ; ceux-ci (notamment ceux liés à la SE de π) seront ensuite affinés lors de l'estimation finale (par une méthode itérative).

2 Cadre géométrique

Notations. Dans les paragraphes qui suivent, les vecteurs $\mathbf{e}_1$, $\mathbf{e}_2$ et $\mathbf{e}_3$ désignent ceux de la base canonique, c.-à-d. vérifiant $(\mathbf{e}_1 \mid \mathbf{e}_2 \mid \mathbf{e}_3) = \mathsf{I}$.

Modèle de caméra. Le paramétrage intrinsèque de la caméra induit une transformation affine A du plan de l'image, dont la matrice (dite *de calibrage*), soit

$$\mathsf{A} = \begin{pmatrix} f & 0 & u_0 \\ & -\tau f & v_0 \\ & & 1 \end{pmatrix},$$

comporte quatre degrés de liberté : la distance focale f, un facteur d'échelle τ et les coordonnées cartésiennes u_0, v_0 du point principal.

Ce modèle de la caméra est dit *modèle de caméra CCD* [4, §5.1], dans lequel les pixels de l'image sont supposés être rectangulaires, τ représentant le rapport des dimensions d'un pixel.

Nous appelons *perspectivité* P la restriction de la projection centrale au plan observé π et notons P sa matrice (d'ordre 3). Par *homographie inter-vues* H_{cj}, nous désignons la composition des perspectivités $P_j \circ P_c^{-1}$ correspondant au déplacement de la caméra d'une configuration clé c vers une configuration j ; la matrice de H_{cj} (d'ordre 3) est notée H_{cj}.

La conique absolue (CA) et son image (ICA). La conique absolue (CA) est une conique imaginaire du plan de l'infini. Elle a une caractéristique remarquable [4, §7.5] : la matrice $\boldsymbol{\omega}$ de son image (ICA) ne dépend que du paramétrage intrinsèque de la caméra. Plus précisément, elle se décompose en $\boldsymbol{\omega} = \mathsf{A}^{-T}\mathsf{A}^{-1}$. Nous notons la matrice de l'ICA duale par $\boldsymbol{\omega}^* = \boldsymbol{\omega}^{-1} = \mathsf{A}\mathsf{A}^T$.

Les points cycliques (PC) et leur images (IPC). Les points cycliques (PC) de π sont deux points imaginaires conjugués de la droite à l'infini, correspondant à l'intersection de π et Ω [4, §1.7.3]. Ils caractérisent la structure euclidienne de π ; en effet, pour *tout* choix d'un repère orthonormé de π, les PC ont une représentation homogène unique (à un facteur près), via les vecteurs isotropiques $\mathbf{i}_\pm = (1, \pm i, 0)^T$. Si l'on note $\mathbf{x}_\pm \sim \mathsf{P}\mathbf{i}_\pm$ les vecteurs

des images des PC (IPC) , étant donné que les PC sont des points fixes de toute similitude du plan, les $\mathbf{x}_\pm$ peuvent être mis sous la forme :

$$\mathbf{x}_\pm \sim \mathsf{Q}\mathbf{i}_\pm = \beta(1,0,\mu)^T \pm i(\alpha,1,\alpha\mu+\lambda)^T, \quad (1)$$

où la matrice Q vérifie :

$$\mathsf{Q} = \mathsf{P}\mathsf{S}^{-1} = \begin{pmatrix} \beta & \alpha & 0 \\ 0 & 1 & 0 \\ \beta\mu & \alpha\mu+\lambda & 1 \end{pmatrix}, \quad (2)$$

pour une certaine matrice S représentant une similitude du plan. Q désigne la composante euclidienne (normalisée) de P, c.-à-d. égale à P à une post-multiplication par S près [4, §1.4.6]. À partir de (2), on peut montrer que la dernière ligne de Q^{-1} correspond au vecteur transposé de la ligne de fuite de π, c.-à-d. à :

$$\mathbf{v}^T = (-\mu,-\lambda,1) = \mathbf{e}_3^T\mathsf{Q}^{-1} \sim \mathbf{e}_3^T\mathsf{P}^{-1}. \quad (3)$$

Les quantités scalaires α, β, λ et μ peuvent être interprétées de la façon suivante. D'une part λ, μ caractérisent la structure affine de π. D'autre part, α, β déterminent la position des IPC sur la ligne de fuite ; ils caractérisent ainsi la «structure angulaire» de la droite de la ligne de fuite de π, une fois celle-ci identifiée.

Conique duale des points cycliques (CDPC). La conique duale des points cycliques (CDPC) est la conique du plan projectif dual, définie tangentiellement par les points cycliques [4, §1.7.3]. La CDPC est une conique dégénérée car elle ne contient que les deux faisceaux de droites passant par les points cycliques. Par conséquent, sa matrice est une matrice (symétrique) de rang 2 de la forme $\mathsf{C}^*_\infty = \mathbf{i}_+\mathbf{i}_-^T + \mathbf{i}_-\mathbf{i}_+^T \sim \mathrm{diag}(1,1,0)$.
Le paramétrage intrinsèque de la caméra transforme la matrice de la CDPC du plan de l'image (par A) en :

$$\mathbf{\Lambda} = \mathsf{A}\mathsf{C}^*_\infty\mathsf{A}^T \sim \mathrm{diag}(1,\tau^2,0). \quad (4)$$

Par conséquent, deux droites sont orthogonales dans l'image ssi elles sont conjugués par rapport à la CDPC transformée par A, c.-à-d. ssi $\mathbf{m}_1^T\mathbf{\Lambda}\mathbf{m}_2 = 0$, où $\mathbf{m}_1$ et $\mathbf{m}_2$ sont les vecteurs de ces droites.

Image de la CDPC (ICDPC). Désignons par Σ l'image (par P) de la CDPC (ICDPC) de π ; Σ a pour matrice :

$$\mathbf{\Sigma} = \mathsf{Q}\mathsf{C}^*_\infty\mathsf{Q}^T \sim \mathsf{P}\mathsf{C}^*_\infty\mathsf{P}^T. \quad (5)$$

Il est à noter que la seule droite qui n'ait pas de point de contact avec l'ICDPC (c.-à-d. pour laquelle il n'existe pas de pôle) est la ligne de fuite de π, ce qui se traduit par $\mathbf{\Sigma}\mathbf{v} = \mathbf{0}$, c.-à-d. $\ker\mathbf{\Sigma} = \{a\mathbf{v}\ ;\ a \in \mathbb{R}^*\}$.
La conique Σ caractérise (de façon duale) la structure euclidienne de π ; elle constitue l'objet de base dans notre méthode d'autocalibrage.

3 Contraintes pour le calibrage plan

Nous présentons ici des contraintes géométriques, induites par la projection centrale des figure planes. Certains résultats sont bien connus (§3.1), d'autres le sont moins (§3.2) ; les points (*ii*) à (*vi*) présentés à la fin de §3.2 sont nouveaux, en l'état actuel de nos connaissances. Les équations (6) sont celles utilisées dans la formulation usuelle du calibrage plan [8][10][11]. L'équation (7) sera celle utilisée comme contrainte dans notre algorithme de calibrage, cf. §4.2.

3.1 La contrainte des points cycliques

Les équations (6) décrivent les relations entre modèle de la caméra et structure euclidienne de π. Elles découlent des propriétés exposées précédemment ; les PC sont situés sur la CA et, par invariance projective, les IPC sont situés sur l'ICA . La mise en équation de cette dernière propriété nous donne en effet $\mathbf{x}_\pm^T\boldsymbol{\omega}\mathbf{x}_\pm = 0$, ou de façon équivalente :

$$\mathbf{x}^T\boldsymbol{\omega}\mathbf{x} = \mathbf{y}^T\boldsymbol{\omega}\mathbf{y}, \quad \mathbf{x}^T\boldsymbol{\omega}\mathbf{y} = 0, \quad (6)$$

où $\mathbf{x}$ et $\mathbf{y}$ sont les parties réelles et imaginaires du vecteur des IPC, soit $\mathbf{x}_\pm$ défini en (3). Les deux équations (6) sont linéaires par rapport aux éléments de $\boldsymbol{\omega}$.

3.2 La contrainte de la ligne principale

Les équations (6) font intervenir la distance focale (via la matrice $\boldsymbol{\omega}$), ce qui peut présenter certains inconvénients, comme expliqué précédemment. Dans ce qui suit, nous proposons une alternative à (6), à savoir une équation décrivant le lieu du point principal comme étant un(e) (segment de) droite. Cette équation exprime la *contrainte de la ligne principale/ligne de centres* [2], dont la représentation est indépendante de la distance focale.
Certains éléments liés à l'orientation du plan observé π peuvent être identifiés dans le plan de l'image ; nous en redéfinissons trois ici (cf. FIG. 1), bien connus des photogrammètres [7]. Le *point nadir* est le point situé à l'intersection du plan de l'image et de la droite passant par le centre optique et perpendiculaire à π ; il peut être aussi défini comme le point de fuite associé à la direction orthogonale à π, à savoir le pôle de la ligne de fuite par rapport à l'ICA duale [4, §.7.6] ; la *ligne principale* (aussi appelée ligne des centres dans [2]) est la droite passant par le point nadir et le point principal ; le *point horizon* est situé à l'intersection de la ligne principale et de la ligne de fuite.

Ligne principale. La droite principale joue un rôle particulier dans le calibrage plan. En effet, il a été montré dans [2] que le lieu du point principal d'une caméra observant une figure plane est un segment de droite, appelé segment principal, situé sur la ligne principale. En d'autres termes, pour tout point pris comme point principal sur ce segment, il existe un plan de la scène sur lequel la figure peut se trouver, tel que la *même* matrice de perspectivité P produit exactement la *même* image de la figure (on comprend mieux en se référant à l'animation disponible à l'URL suivante : `www.irit.fr/~Pierre.Gurdjos/ECCV2002/`).

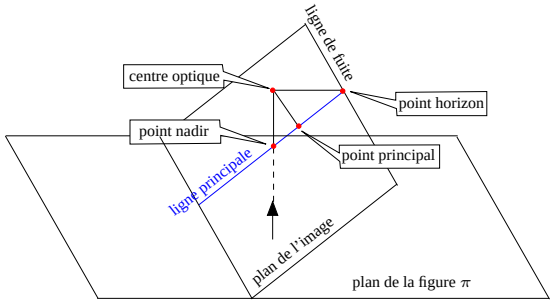

FIG. 1 – Point nadir, point horizon et ligne principale.

Cette propriété remarquable a été employée comme contrainte de calibrage (cf. la contrainte de la ligne des centres dans [2]) ; elle a l'avantage d'être indépendante de la distance focale. Dans un travail similaire [1], il a été aussi montré le point suivant.
(*i*) Le vecteur de la ligne principale peut s'écrire sous une forme indépendante des paramètres intrinsèques, uniquement en fonction de la perspectivité P :

$$\boldsymbol{\phi} = (-\varphi_1, -\varphi_2, \varphi_3 + \varphi_4)^T,$$

où φ_i est une fonction bilinéaire des éléments de P.
Pour être plus précis, cette propriété n'est vérifiée qu'à condition que les pixels soient carrés (c.-à-d. pour $\tau = 1$). Une de nos contributions est de montrer que, si les pixels sont rectangulaires alors le lieu de la ligne principale est un faisceau de droites dont le sommet est le point horizon (ce sommet ne dépend que de la perspectivité P). Une représentation matricielle de la contrainte de la ligne des centres peut être alors donnée ; c'est celle-ci que nous utiliserons dans notre formulation du problème de l'autocalibrage plan (cf. §4.2).

Le faisceau des lignes principales (cf. FIG. 2). Les propriétés suivantes sont établies en présence d'un facteur d'échelle τ arbitraire (les preuves sont omises).
(*ii*) Le point horizon est le pôle de la droite de l'infini par rapport à l'ICDPC. Il a pour vecteur $\boldsymbol{\Sigma}\mathbf{e}_3$, où $\boldsymbol{\Sigma}$ est la matrice donnée en (5), qui ne dépend donc que de Q.
Le point nadir est le pôle de la ligne de fuite par rapport à l'image de la conique absolue duale ; il a pour vecteur $\boldsymbol{\omega}^*\mathbf{v}$ [4, §.7.6, (7.14)], où $\boldsymbol{\omega}^* = \boldsymbol{\omega}^{-1}$ et $\mathbf{v}$ est le vecteur de la ligne de fuite donné en (3).
(*iii*) Le point nadir, le point principal et le point horizon sont des points alignés (cf. FIG. 2), ce qui implique :

$$\det\left(\boldsymbol{\Sigma}\mathbf{e}_3 \mid \mathbf{p}_0 \mid \boldsymbol{\omega}^*\mathbf{v}\right) = 0,$$

où $\mathbf{p}_0 = (u_0, v_0, 1)^T$ est le vecteur du point principal. Du fait que :

$$\mathbf{p}_0 \times \boldsymbol{\omega}^*\mathbf{v} = f^2(\mathbf{p}_0 \times \boldsymbol{\Lambda}\mathbf{v}),$$

où $\boldsymbol{\Lambda}$ est la matrice définie en (4), on en déduit la contrainte suivante :

$$\boxed{\det\left(\boldsymbol{\Sigma}\mathbf{e}_3 \mid \mathbf{p}_0 \mid \boldsymbol{\Lambda}\mathbf{v}\right) = 0.} \qquad (7)$$

L'équation (7) mérite une interprétation. Le vecteur $\boldsymbol{\Lambda}\mathbf{v}$ est celui du point à l'infini de toute droite orthogonale à la ligne de fuite ; l'équation (7) nous dit donc que le point principal se trouve sur la droite orthogonale à la ligne de fuite et passant par le point nadir, à savoir la ligne principale. Par conséquent, nous venons de donner une nouvelle représentation matricielle de la contrainte de la ligne principale/des centres. Elle met en relation le facteur d'échelle (via $\boldsymbol{\Lambda}$), le point principal et les quatre paramètres liés à la structure euclidienne de π (via $\boldsymbol{\Sigma}$).
De plus,
(*iv*) le vecteur de la ligne principale ne dépend que de la matrice Q et du facteur d'échelle τ ; il est de la forme :

$$\boldsymbol{\phi} = (0, -\varphi_2, \varphi_4)^T + \tau^2\,(-\varphi_1, 0, \varphi_3)^T \sim \mathbf{d}_1 + \tau^2\mathbf{d}_2, \quad (8)$$

où $\mathbf{d}_1 = (\boldsymbol{\Sigma}\mathbf{e}_3) \times \mu\mathbf{e}_1$ et $\mathbf{d}_2 = (\boldsymbol{\Sigma}\mathbf{e}_3) \times \lambda\mathbf{e}_2$.
Nous savons que si d_1, d_2 sont deux droites de vecteurs $\mathbf{d}_1, \mathbf{d}_2$ alors toute droite passant par le point d'intersection de d_1 et d_2 peut être représentée par l'équation

$$x_1\mathbf{d}_1 + x_2\mathbf{d}_2 = 0, \qquad (9)$$

où x_1, x_2 sont deux scalaires non nuls simultanément.
(*v*) L'équation (9) représente donc un système linéaire de droites, à savoir un certain faisceau de droites $\mathscr{F}$, dont le sommet coïncide avec le point horizon ; en effet, on vérifie aisément que ce sommet a pour vecteur :

$$\mathbf{d}_1 \times \mathbf{d}_2 \sim \boldsymbol{\Sigma}\mathbf{e}_3. \qquad (10)$$

(*vi*) La ligne principale est paramétrée (projectivement) par $\frac{x_2}{x_1} = \tau^2$, relativement au faisceau de droites $\mathscr{F}$.

(*vii*) Le lieu de la ligne principal est donc un faisceau de droites (cf. FIG 2), de sommet le point horizon, qui admet la représentation (8) ; celle-ci ne dépend que de Q, où Q est la matrice donnée en (2).

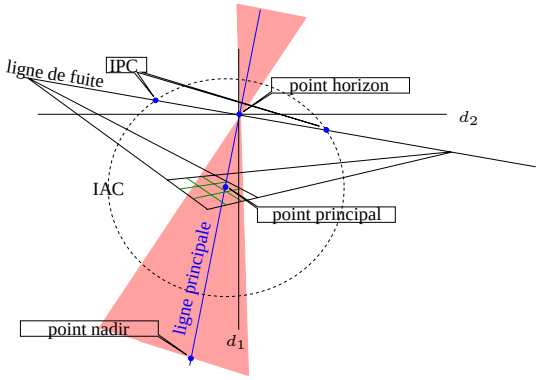

FIG. 2 – Lorsque τ est inconnu, le lieu de la ligne principale est un faisceau de droites dont le sommet est le point horizon.

4 Autocalibrage plan

Nous supposons que nous avons réalisé l'acquisition de n vues d'une figure plane. Le problème consiste à estimer les p ($0 < p \leq n+3$) paramètres intrinsèques (soient, au plus,

u_0, v_0, τ et les f_j avec $1 \leq j \leq n$) et les $q \leq 4$ paramètres liés à la structure euclidienne du plan observé π (soient, au plus, α, β, λ et μ), associés à une vue-clef.
Tout d'abord, plaçons-nous dans le cas où $q = 0$, c.-à-d. lorsqu'une représentation euclidienne de la figure peut être donnée. Nous pouvons donc estimer les n matrices des perspectivités et disposons donc d'un système de $2n$ équations (6), en utilisant (1), ou n équations (7) linéaires, en utilisant (3,5). Le problème est alors celui du *calibrage plan* qui a une solution linéaire par rapport aux éléments de $\boldsymbol{\omega}_j$ [8][11], même dans le cas où f_j n'est pas constante [2][8]. D'un point de vue géométrique, ce problème peut être envisagé comme celui de l'ajustement des images de la CA aux n images des PC ou de celui l'intersection de n lignes principales [2].
Nous considérons maintenant le cas général de l'autocalibrage, c.-à-d. lorsque la structure euclidienne de π est inconnue. Dans ce cas, le « point de départ » est la structure projective de π, qui est accessible dès lors que les (matrices des) homographies inter-vues peuvent être estimées. B. Triggs a proposé dans [10] une solution à ce problème

4.1 Formulations non linéaires existantes

Approche de B. Triggs [10]. Par rapport au problème du calibrage plan, l'« astuce » proposée par B. Triggs dans [10] est d'introduire $q = 4$ inconnues supplémentaires afin de représenter les IPC dans une certaine vue, appelée vue-clef ; les IPC de la vue-clef sont ensuite transférées dans les autres images, via les homographies inter-vues. En reprenant les équations (6) nous disposons, pour chaque vue, de $2n$ équations non linéaires de la forme[1] :

$$(\mathsf{H}_{cj}\mathbf{x}_{\pm})^T \boldsymbol{\omega}_j (\mathsf{H}_{cj}\mathbf{x}_{\pm}) = 0, \qquad (11)$$

où H_{cj} est la matrice inter-vues, de la vue-clef vers la $j^{\text{ième}}$ vue ($1 \leq j \leq n$) et $\boldsymbol{\omega}_j$ la matrice de l'IAC dans la $j^{\text{ième}}$ vue.
Au total, étant données les n matrices H_{cj} des homographies inter-vues ($\mathsf{H}_{cc} = \mathsf{I}$), cette méthode d'autocalibrage plan consiste à résoudre (au sens des moindres carrés) un système de $2n$ équations non linéaires (11), pour r inconnues ($7 < r \leq n + 7$) . Si $r = 8$, au moins 4 vues sont nécessaires ; si $r = n + 7$, au moins 7 le sont.
B. Triggs a détaillé, dans [10], une méthode itérative permettant d'obtenir une solution dans le cas de paramètres intrinsèques constants ($r = 8$) ; celle-ci a été ensuite généralisée dans [4, § 18.7] au cas d'une distance focale spécifique («variable») à chaque vue ($r = n+7$). Ces méthodes présentent néanmoins un inconvénient qui est mentionné : la nécessité de disposer de valeurs initiales pour les inconnues (du moins pour la/les distance(s) focale(s)) qui soient proches des solutions. Un autre inconvénient majeur apparaît aussi dans le cas où le nombre d'inconnues est maximum (c.-à-d. pour n distances focales variables) : si n croît, alors la sensibilité de l'algorithme à l'initialisation des inconnues ainsi que le temps de calcul (avec une complexité algorithmique de $\mathcal{O}(r^3)$ par itération) augmentent de façon significative.

Approche d'E. Malis [5]. Parmi les formulations non linéaires existantes, il est nécessaire de citer les travaux de E. Malis [5]. Afin de réduire le bruit sur les matrices inter-vues, cet auteur suggère tout d'abord d'imposer, selon une approche par « factorisation », une contrainte de rang sur la matrice bloc, dont les éléments (k, l) sont les homographies inter-vues H_{lk} ($1 \leq k, l \leq n$). Cette idée n'est pas nouvelle, mais sa mise en œuvre ici diffère de celle proposée par B. Triggs dans une version longue de [10]. En fait, cette « factorisation » des données peut être utilisée, indépendamment de la méthode d'autocalibrage choisie ; malheureusement, dans les tests que nous avons effectués, nous n'avons pas été convaincus par l'apport de cette étape préliminaire (coûteuse en temps de calcul) quant à un possible gain de précision. Pour l'autocalibrage plan, proprement dit, Malis a ensuite mis en évidence une propriété algébrique pour toute matrice d'homographie inter-vues «calibrée», c.-à-d. pour $\tilde{\mathsf{H}}_{jk} = \mathsf{A}^{-1}\mathsf{H}_{jk}\mathsf{A}$, de la forme :

$$\mathrm{svd}(\tilde{\mathsf{H}}_{jk} [\mathbf{n}_j]_{\times}) = \mathrm{diag}(\sigma, \sigma, 0), \qquad (12)$$

où $\mathbf{n}_j$ est le vecteur normal au plan observé, exprimé dans le repère de la $j^{\text{ième}}$ caméra, $[\mathbf{n}_j]_{\times}$ est la matrice antisymétrique usuelle associée à $\mathbf{n}_j$ ($[\mathbf{n}_j]_{\times}\, \mathbf{n}_j = \mathbf{0}$) et $\mathrm{svd}(\mathsf{X})$ est une fonction retournant la matrice diagonale des valeurs singulières de X. La propriété (12) exprime le fait que la matrice $\tilde{\mathsf{H}}_{jk} [\mathbf{n}_j]_{\times}$ a deux valeurs singulières égales à σ (et une nulle) et équivaut à deux contraintes sur les paramètres intrinsèques ; le critère de minimisation proposé par E. Malis est basé sur celle-ci.
Au total, la contribution proposée par Malis semble très intéressante, même s'il s'agit d'un algorithme multi-étapes, opérant à partir de « factorisations » successives, qui reste délicat à mettre en œuvre et dont la convergence reste difficile à prouver. Pour cette raison et celle évoquée plus haut, nous n'avons pas comparé les résultats d'autocalibrage obtenus par notre algorithme avec ceux de l'algorithme de Malis.

4.2 Formulation non linéaire proposée

La contrainte de la ligne principale, donnée en §3.2, Éq. (7), est exprimée à partir de la composante euclidienne de la perspectivité P, représentée par la matrice Q. Nous supposons que le vecteur $\mathbf{v}$ et la matrice $\boldsymbol{\Sigma}$ représentent respectivement la ligne de fuite et l'IDCP dans la vue-clef.
Nous devons adapter la contrainte (7) afin de pouvoir y introduire les matrices des homographies inter-vues. Cela se fait très facilement puisque, par H_{cj}, le vecteur $\mathbf{v}$ se transforme en $\mathbf{v}_j = \mathsf{H}_{jc}^{-T}\mathbf{v}$ et la matrice $\boldsymbol{\Sigma}$ en $\boldsymbol{\Sigma}_j = \mathsf{H}_{jc}\boldsymbol{\Sigma}\mathsf{H}_{jc}^T$. Il s'ensuit que la contrainte de la ligne principale, pour la $j^{\text{ième}}$ vue, s'écrit :

$$\det\left(\mathbf{p}_0 \mid \boldsymbol{\Lambda}\mathsf{H}_{jc}^{-T}\mathbf{v} \mid \mathsf{H}_{jc}\boldsymbol{\Sigma}\mathsf{H}_{jc}^T\mathbf{e}_3\right) = 0. \qquad (13)$$

[1] (11) est la forme «condensée» de :

$$\mathbf{u}_j^{\top}\boldsymbol{\omega}_j\mathbf{u}_j = \mathbf{v}_j^{\top}\boldsymbol{\omega}_j\mathbf{v}_j, \quad \mathbf{u}_j^{\top}\boldsymbol{\omega}_j\mathbf{v}_j = 0,$$

où $(\mathbf{u}_j \mid \mathbf{v}_j) \sim \mathsf{H}_{cj}\,(\mathbf{x}_j \mid \mathbf{y}_j)$.

Rappelons que notre formulation a un nombre réduit d'inconnues, indépendant du nombre de vues (les distances focales n'interviennent pas). Ce sont les p_c ($0 \leq p_c \leq 3$) paramètres intrinsèques constants, soient u_0, v_0 et τ, et les 4 paramètres $\alpha, \beta, \lambda, \mu$ de $\mathbf{\Sigma}$ (avec $\mathbf{v} = (-\mu, -\lambda, 1)^T$), liés à la structure euclidienne de π. Le problème de l'autocalibrage se formule donc comme suit.

Problème. Étant données les n matrices H_{cj} des homographies inter-vues ($\mathsf{H}_{cc} = \mathsf{I}$), il s'agit de résoudre (au sens des moindres carrés) un système de n équations non linéaires (13), pour r paramètres ($4 \leq r \leq 4 + p_c$). Une fois ces paramètres estimés, les différentes distances focales sont calculées par une méthode similaire à celle décrite dans [1, §6.2] ou dans [8][11]. Par conséquent, Si $p_c = 3$ ($r = 7$), au moins 7 vues sont nécessaires. Si $p_c = 0$, le problème résolu est celui de la *rectification euclidienne* de l'image et la solution proposée généralise celle décrite dans [4, §1.7.5].

Rectification euclidienne. La rectification euclidienne de l'image consiste à appliquer la transformation inverse Q^{-1} aux points de l'image, où Q est la perspectivité définie dans (2), afin d'obtenir une reconstruction euclidienne de la figure (et du plan π), à une similitude près. Il est donc possible de dissocier (à partir de l'hypothèse courante que $u_0 = v_0 = 0$ et $\tau = 1$, après recentrage des données) l'estimation des 4 paramètres $\alpha, \beta, \lambda, \mu$ liés à la perspectivité Q de celle des paramètres intrinsèques (ici restreints aux distances focales). Ce résultat nous paraît assez intéressant pour être mentionné. Il est illustré en §5, cf. FIG. 5.

Solution(s). Dans l'équation (13), nous définissons la matrice $\mathsf{M}_j = \left(\mathbf{p}_0 \mid \mathbf{\Lambda}\mathsf{H}_{jc}^{-T}\mathbf{v} \mid \mathsf{H}_{jc}\mathbf{\Sigma}\mathsf{H}_{jc}^{T}\mathbf{e}_3\right)$.
Une première solution au problème ci-dessus peut être obtenue par minimisation du critère $\mathcal{C} = \sum_{j=1}^{n} (\det \mathsf{M}_j)^2$. Un des aspects intéressants de notre formulation est qu'il est facile de calculer analytiquement le gradient de $\det \mathsf{M}_j$. Nous ne donnons pas plus de détails mais rappelons les propriétés matricielles suivantes (nous omettons les indices pour M_j et H_{jc}). D'une part :

$$\det \mathsf{M} = (H_{31} + \mu H_{33}) \det \mathsf{N}_1 + (H_{32} + \lambda H_{33}) \det \mathsf{N}_2,$$

où, pour $k = 1, 2$,

$$\mathsf{N}_k = \left(\ \mathsf{H}\mathbf{\Sigma}(:,k) \mid \mathsf{M}(:,2) \mid \mathsf{M}(:,3) \ \right) \ ;$$

$\mathsf{X}(:,c)$ désigne la $c^{\text{ième}}$ colonne d'une matrice X et $X(l,c)$ son élément (l,c). D'autre part,

$$\frac{d \det \mathsf{M}}{dx} = \operatorname{trace}\left\{ \frac{d\mathsf{M}}{dx} \operatorname{adj} \mathsf{M} \right\}, \qquad (14)$$

où $\operatorname{adj} \mathsf{M} = (\det \mathsf{M})\, \mathsf{M}^{-1}$ désigne la matrice adjointe de M. Le gradient de $\det \mathsf{M}_j$ est calculé à partir de (14) pour $x = \alpha, \beta, \lambda, \mu, u_0, v_0, \tau$.
Une deuxième méthode de résolution consiste à donner une interprétation géométrique à la quantité $\det \mathsf{M}$. En effet, nous savons que $\det \mathsf{M} = \mathsf{M}(:,1)^T \left(\mathsf{M}(:,2) \times \mathsf{M}(:,3)\right)$ (produit mixte des colonnes de M) ; ceci veut dire que la quantité :

$$\delta(\mathsf{M}) = \frac{\left|\mathsf{M}(:,1)^T \left(\mathsf{M}(:,2) \times \mathsf{M}(:,3)\right)\right|}{\sqrt{\left(\mathsf{M}(:,2) \times \mathsf{M}(:,3)\right)^T \mathbf{\Lambda} \left(\mathsf{M}(:,2) \times \mathsf{M}(:,3)\right)}} \qquad (15)$$

représente la distance du point, de vecteur $\mathsf{M}(:,1)$, à la droite, à savoir la ligne principale (cf. §3.2), de vecteur $\mathsf{M}(:,2) \times \mathsf{M}(:,3)$. La solution est alors obtenue par minimisation du critère $\mathcal{C} = \sum_{j=1}^{n} (\delta(\mathsf{M}_j))^2$.
Cette approche est celle utilisée dans la version actuelle de notre algorithme En ce qui concerne le codage proprement dit, nous avons utilisé la fonction `lsqnonlin` (moindres carrés non linéaires), disponible dans la bibliothèque mathématique spécialisée *MATLAB Optimization Toolbox* [6], avec itérations du type «Levenberg-Marquardt».
L'algorithme général est récapitulé en TAB. 1.

4.3 Solutions exactes

Nous donnons ici des solutions exactes, à partir d'une l'hypothèse dite de «fronto-parallélisme», c.-à-d. lorsque, pour la vue-clef, la figure se situe sur un plan fronto-parallèle. Comme il a été mentionné précédemment, il s'ensuit que la ligne de fuite est la droite de l'infini, c.-à-d. nous avons $\lambda = \mu = 0$ dans (3). Bien sûr, ceci simplifie grandement le problème de l'autocalibrage. Nous rappelons que cette hypothèse de fronto-parallélisme n'est utilisée que dans une phase préliminaire de notre algorithme (cf. TAB. 1), afin de pouvoir assigner des valeurs initiales aux paramètres à estimer. Lors de la résolution itérative (finale), cette hypothèse est levée. Vis-à-vis du cadre géométrique détaillé en §2, ce qui change ici est que, dans la vue-clef, les IPC sont situés, à la fois, sur la droite de l'infini et sur l'IAC ; il est facile de vérifier que les seuls points vérifiant ces propriétés sont les points imaginaires conjugués ayant pour vecteurs :

$$\left(\tau^{-1}, 0, 0\right)^T \pm i\,(0, 1, 0)^T \sim (1, \pm\tau i, 0)^T . \qquad (16)$$

Par conséquent, en se référant en (1), on en déduit que $\alpha = 0$ et donc qu'il est possible de «retrouver» la structure euclidienne de π au paramètre $\beta = \tau^{-1}$ près. Nous imaginons maintenant plusieurs scénarios pour lesquels des solutions exactes peuvent être calculées, en l'occurrence les cas où le facteur d'échelle τ est connu/inconnu et la distance focale f est constante/variable.
Lorsque τ est connu, l'hypothèse de fronto-parallélisme implique que $\beta = \tau^{-1}$ l'est aussi et le problème de l'autocalibrage plan devient celui du calibrage plan, pour lequel – comme mentionné précédemment – des solutions bien connues existent [8, 11, 1]. Intéressons-nous donc au cas où τ est inconnu.
De (16), nous en déduisons que, dans la vue-clef, les IPC – et donc la SE de π – sont paramétrés par une seule quantité τ. Les homographies inter-vues pemettent de transférer ces IPC dans les autres vues, à partir du seul τ. Soit H_{cj} la matrice de homographie inter-vues, de la vue-clef vers la $j^{\text{ième}}$

Algorithme LIGPAL	
Données :	n matrices homographies inter-image H_{cj} ($1 \leq j \leq n$, $\mathsf{H}_{cc} = \mathsf{I}$).
Normalisations :	Soit la $r \times s$ la résolution de l'image ; soit $\mathsf{T} = \begin{pmatrix} \max(r,s) & 0 & r/2 \\ 0 & \max(r,s) & s/2 \\ 0 & 0 & 1 \end{pmatrix}$. Calculer $\bar{\mathsf{H}}_{cj} = \mathsf{T}^{-1}\mathsf{H}_{cj}\mathsf{T}$, puis normaliser $\bar{\mathsf{H}}_{cj}$ de telle façon que $\left\|\bar{\mathsf{H}}_{cj}\right\|_F = 1$.
Inconnues (7) :	$\alpha, \beta, \lambda, \mu$ via la matrice Σ (avec $\mathbf{v} \in \ker(\mathbf{\Sigma})$) ; u_0, v_0 via $\mathbf{p}_0 = (u_0, v_0, 1)^T$ et τ via $\mathbf{\Lambda}$
Résolution	Soit $\mathsf{M}_{jc} = \left(\mathbf{p}_0 \mid \mathbf{\Lambda}\mathsf{H}_{jc}^{-T}\mathbf{v} \mid \mathsf{H}_{jc}\mathbf{\Sigma}\mathsf{H}_{jc}^{T}\mathbf{e}_3\right)$.
(Moindres carrés non-linéaires) :	Chercher $\arg\min_{\substack{\alpha,\beta,\lambda,\mu \\ u_0,v_0,\tau}} \sum_{j=1}^{n} (\delta(\mathsf{M}_{jc}))^2$, où $\delta(\mathsf{M}_{jc})$ est défini en (15). Calculer les n distances focales f_j selon la procédure décrite dans [1, §6.2].

TAB. 1 – Algorithme LIGPAL .

vue. Pour cette vue, les équations (11) sont alors équivalentes à

$$\mathbf{h}_1^T \boldsymbol{\omega}_j \mathbf{h}_1 - \tau^2 \mathbf{h}_2^T \boldsymbol{\omega}_j \mathbf{h}_2 = 0, \quad \mathbf{h}_1^T \boldsymbol{\omega}_j \mathbf{h}_2 = 0, \qquad (17)$$

où $\mathbf{h}_k$ est la $k^{\text{ième}}$ colonne de H_{cj}. Il est à noter que la seconde équation de (17) est linéaire pour les 4 éléments de $\boldsymbol{\omega}_j$ et que, par conséquent, une solution peut être obtenue à partir de 5 vues. Quant à la première équation de (17), en utilisant le fait que $\tau^2 = \omega_{11}/\omega_{22}$, où ω_{kl} est l'élément (k, l) de $\boldsymbol{\omega}_j$, nous pouvons la réécrire ainsi :

$$\omega_{22} h_1^T \boldsymbol{\omega}_j \mathbf{h}_1 - \omega_{11} \mathbf{h}_2^T \boldsymbol{\omega}_j \mathbf{h}_2 = 0.$$

C'est une équation polynomiale de degré 2, à plusieurs variables (à savoir les éléments de $\boldsymbol{\omega}_j$), avec uniquement ω_{11} and ω_{22} élevés au carré. Bien qu'il existe des schémas numériques de résolution adaptés à ce type d'équation, il nous semble utile de s'intéresser à des cas particuliers, que nous décrivons sans donner trop de détails, à cause du manque de place. Nous notons que, dans la vue-clef, l'hypothèse de fronto-parallélisme rend impossible le calcul de f [8] ; par conséquent, nous l'ignorerons en écrivant ($\boldsymbol{\omega} = \boldsymbol{\omega}_j$).

Point principal connu. Deux vues, celle « fronto-parallèle » et tout autre vue suffisent pour autocalibrer ; les seules inconnues sont le facteur d'échelle τ et la distance focale f (éventuellement différente) dans cette vue. Regardons à nouveau les équations (17). Il est toujours possible d'effectuer un changement de repère de telle façon que $u_0 = v_0 = 0$, ce qui implique alors que la matrice de l'IAC s'écrive $\boldsymbol{\omega} = \mathrm{diag}(\tau^2, 1, \tau^2 f^2)$. Afin de simplifier, remplaçons ces inconnues par $a = \tau^2$ et $b = \tau^2 f^2$. Les équations (17) deviennent alors :

$$H_{21}^2 + a(H_{11}^2 - H_{22}^2) + bH_{31}^2 - abH_{32}^2 - a^2 H_{12}^2 = 0$$
$$H_{21}H_{22} + aH_{11}H_{12} + bH_{31}H_{32} = 0.$$

Résoudre la seconde équation pour a est trivial. En substituant la solution dans la première équation, nous obtenons une équation du second degré en b. Donner les solutions exactes pour τ and f, à partir des éléments de H_{cj}, ne pose aucune difficulté particulière.

Point principal inconnu et distance focale constante. À partir de trois vues (pour quatre inconnues f, τ, u_0, v_0), nous procédons de la même façon et obtenons une équation de degré 3 pour une seule variable, pour laquelle des solutions exactes peuvent être calculées.

Point principal inconnu et distance focale variable. À partir de quatre vues, nous procédons de la même façon et obtenons une équation de degré 4 pour une seule variable, pour laquelle des solutions exactes peuvent être aussi calculées.

5 Validation expérimentale

Dans [3], des simulations ont été réalisées afin d'évaluer les performances de notre algorithme (LIGPAL, décrit en TAB. 1), en le comparant à celles obtenues par celui de B. Triggs (cf §4.1), que nous appellerons PTCIRC. Il en ressort que, pour des variations de focales de moyenne amplitude, un assez grand nombre d'images ($n \approx 15$) est nécessaire pour que la précision de l'algorithme LIGPAL soit identique à celle de PTCIRC, en insistant sur le fait que LIGPAL a une complexité algorithmique nettement moindre. Cette nécessité d'un assez grand nombre d'images s'explique par le fait que les mouvements critiques pour LIGPAL constitue vraisemblablement un sur-ensemble de ceux pour PTCIRC mais, qu'à partir de $n \geq 15$, l'incidence des mouvements critiques spécifiques à LIGPAL sur les résultats devient négligeable.

Description de la simulation. Dans ce travail, nous mettons en évidence une expérience simulant de larges variations de la distance focale, vis-à-vis de laquelle l'algorithme LIGPAL obtient les meilleurs résultats, même pour un faible nombre de vues. Chaque expérience décrite ci-dessous a été réalisée 500 fois, pour un nombre de vues n variant de 7 à 15. Nous décrivons maintenant le protocole pour chaque test unitaire. Lorsque rien n'est mentionné, le choix (ou tirage) aléatoire d'une valeur obéit à une loi aléatoire uniforme. La distance focale de la caméra est différente pour chaque vue : sa valeur (en pixels) est tirée aléatoirement dans l'intervalle $[800; 3600]$; la ré-

solution des images est de 640×480 pixels. Les autres paramètres intrinsèques sont constants dans les n vues : la position (u_0, v_0) du point principal est choisie dans un cercle, de même centre que l'image et de rayon 100 pixels, et le facteur d'échelle dans l'intervalle $[0.9; 1.1]$. Les différents prises de vues correspondent à différentes orientations/positions de la caméra, choisies de la façon décrite ci-après ; la notation $v \sim \mathcal{N}(\bar{v} \pm \sigma_v)$ indique que la variable v suit une loi normale de moyenne $\bar{v}$ et d'écart type σ_v. Un nuage de 100 points est généré aléatoirement dans la vue-clef puis est transféré dans les autres vues ; les coordonnées (en pixels) des points sont pertubées avec un bruit $b \sim \mathcal{N}(0 \pm 1)$. Pour toute vue, l'angle entre le plan de l'image et le plan observé (angle d'ascension) est $\theta_1 \sim \mathcal{N}(30° \pm 20°)$ et l'angle d'azimuth est $\theta_2 \sim \mathcal{N}(0° \pm 45°)$, à l'exception de la première vue pour laquelle $\theta_1 \sim \mathcal{N}(0° \pm 10°)$. Une rotation d'angle $\theta_3 \sim \mathcal{N}(0° \pm 180°)$ autour de l'axe optique est finalement effectuée. La caméra vise approximativement un certain point du plan ; la distance entre le centre optique et ce point est choisie de telle façon que le nuage de points occupe une surface a peu près constante dans toutes les vues.

Nous donnons maintenant quelques précisions sur la mise en œuvre des algorithmes (cf. TAB. 1). Les coordonnées des points ont été normalisées de telle façon qu'elles soient comprises entre -1 et 1. Les matrices des homographies inter-vues H sont estimées, à partir des coordonnées bruitées des points, par l'algorithme *DLT* (*Direct Linear Transformation*) décrit dans [4, §3.1], et sont normalisées de telle façon que $\|\mathsf{H}\|_F = 1$. Quant à l'initialisation des distance focales pour PTCIRC, nous utilisons une valeur correspondant à 1280 pixels (c.-à-d. deux fois la largeur de l'image). Nous avons calculé les erreurs absolues sur les coordonnées (u_0, v_0) du point principal ainsi que les erreurs relatives sur le rapport d'échelle κ et sur les différentes focales f_j, $1 < j \leq n$. Dans le calcul des erreurs relatives sur les f_j, nous avons rejeté des valeurs aberrantes, c.-à-d. hors de l'intervalle $]400; 5400[$ (soit une erreur relative de 50% sur chacune des bornes de l'intervalle initial) ; ceci doit être pris en compte dans la lecture des résultats obtenus. Les taux de rejets sont présentés en TAB. 2 ; nous notons que la méthode LIGPAL a les plus faibles taux de rejets. Les moyennes des erreurs obtenues sont affichées dans la figure FIG. 3. Pour $n = 11$ vues, les erreurs sur les coordonnées du point principal restent assez élévées (aux environs de 15 pixels), le facteur d'échelle est, par contre, estimé avec une très bonne précision (à peu près 0.5% d'erreur relative dans les deux cas) ; la focale est estimée plus précisément avec LIGPAL qu'avec PTCIRC (5% contre 7.5%).

N. vues	7	9	11	13	15
LIGPAL	2.6%	2.75%	0.94%	0.13%	0.11%
PTCIRC	4.9%	2.05%	1.16%	1%	0.96%

TAB. 2 – Taux de rejets en fonction du nombre de vues.

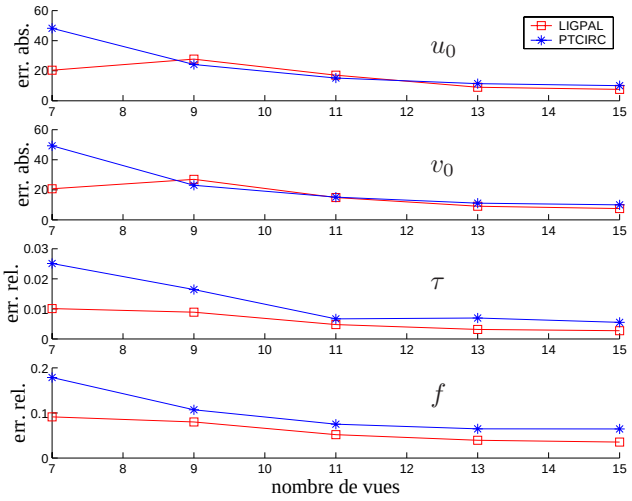

FIG. 3 – Comparaison des résultats obtenus par les algorithmes LIGPAL (proposée) et PTCIRC [10].

Comparaison des algorithmes à partir d'images réelles. Nous comparons ici les résultats de calibrage obtenus par les méthodes LIGPAL et PTCIRC, à partir d'images réelles. Nous souhaitons utiliser comme valeurs de référence les paramètres estimés par la méthode linéaire de calibrage plan [8][11], que nous appelons PTCIRC-LIN. Pour ce faire, nous utilisons 15 vues d'une mire plane, sous la forme d'un damier, acquises par un appareil numérique *Nikon Coolpix 800* avec une résolution 640×480 (cf. FIG. 4). Les 15 matrices de perspectivités (utilisées par PTCIRC-LIN) et les 14 matrices des homographies inter-vues (utilisées par LIGPAL et PTCIRC) ont été calculées à partir des coins des carrés noirs.

Après calibrage et autocalibrage, les valeurs estimées pour le point principal et le facteur d'échelle sont : $(308, 250, 1.008)$ pour PTCIRC-LIN, $(325, 253, 0.999)$ pour PTCIRC et $(325, 260, 0.999)$ pour notre algorithme LIGPAL. Il est à noter que les différences sur le facteur d'échelle ne dépassent pas 1%. Nous supposons que la mise au point automatique entraînent des variations de la distance focale ; c'est la raison pour laquelle nous considérons celle-ci différente dans chaque vue. Les différentes distances focales calculées par l'algorithme PTCIRC sont : 1368, 1390, 1383, 1352, 1357, 1357, 1371, 1322, 1346, 1352, 1358, 1345, 1390, 1394, 1387, de moyenne 1364 et écart-type 20 (c.-à-d. 1%). Pour LIGPAL, la table 3 donne les différences relatives (%) avec PTCIRC-LIN . En parenthèses, les différences relatives avec PTCIRC sont affichées (une valeur négative veut dire « LIGPAL est plus proche de PTCIRC-LIN que PTCIRC »). Les différences entre les distances focales, estimées par les algorithmes de calibrage et d'autocalibrage, sont très faibles (dans la plupart des cas, inférieures à 1%). Enfin, la première image a été rectifiée, à partir de l'estimation des 4 paramètres liés à la structure euclidienne de π. L'image qui en résulte (cf. Fig. 4) semble correcte.

FIG. 5 – Rectification automatique d'une figure plane « inconnue », en supposant le point principal et le facteur d'échelle connus.

0.5 (−0.1)	0.4 (−0.1)	1.2 (−0.1)	0.9 (−0.1)	0.7 (0.1)
0.0 (−0.1)	0.3 (0.1)	1.6 (1.0)	0.2 (0.1)	0.3 (0.2)
0.6 (0.3)	0.4 (0.2)	0.9 (0.1)	1.3 (0.0)	0.1 (−0.1)

TAB. 3 – Comparaison des valeurs obtenues pour les distances focales par l'algorithme LIGPAL avec celles obtenues par un algorithme de calibrage plan, à partir des 15 vues de la figure 4.

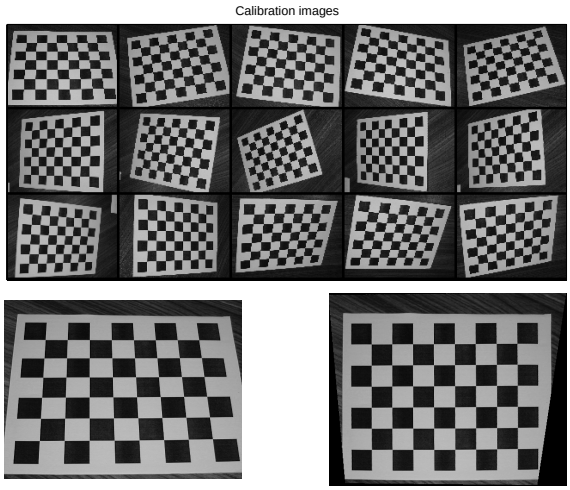

FIG. 4 – En haut et à gauche : images originales ; à droite : image rectifiée après autocalibrage plan.

Vidéos d'une bande dessinée. Nous avons filmé plusieurs séquences vidéo d'une bande dessinée ouverte et posée à plat, à partir d'un caméscope numérique (*Panasonic NV-MX 300*). Les acquisitions ont débuté avec la bande dessinée située approximativement dans un plan frontoparallèle. Les vidéos ont été traitées automatiquement ; dans chaque séquence, nous n'avons utilisé qu'une vue sur dix ; la figure 6 montre 8 parmi les 20 vues utilisées pour une de ces séquences. Des point d'intérêts ont été extraits afin d'estimer les homographies inter-vues (en utilisant une méthode de type RANSAC). Nous avons calculé les solutions exactes données en §4.3, dans le cas d'un point principal supposé être situé au centre de l'image. Les distances focales ont été calculées pour toutes les vues, sauf la première, et les facteurs d'échelle pour toutes les vues. Une seule valeur pour τ a été finalement retenue, en utilisant des outils issus des statistiques robustes.

À partir de ces estimations, une première approximation de la structure euclidienne du plan observé a pu être estimée ; à partir de celle-ci et des paramètres intrinsèques, les différentes poses des caméras peuvent être calculées [9].

Toutes ces estimations ont été utilisées pour initialiser une procédure d'« ajustement de faisceaux » (estimant aussi la position des points sur le plan 3D et la distortion radiale). Une solution a été obtenue au bout de 2 ou 3 itérations, chacune de celles-ci nécessitant quelques secondes de temps de calcul.

Ces résultats ont été comparé avec ceux obtenus en calibrant le caméscope, après avoir filmé une vidéo d'une mire de calibrage. Le facteur d'échelle a été estimé avec une erreur relative de 0.2% (soit 1.0893, à comparer à la valeur 1.0919 issue du calibrage). Le point principal a été estimé avec une précision d'environ 10 pixels. Quant aux différentes focales, pour la première moitié de la séquence, les résultats ont été mauvais, comme nous pouvions nous y attendre (le plan de l'image et le plan observé étaient presque parallèles). Dans la seconde moitié, la valeur moyenne des estimations est 1366, ce qui indique une erreur relative de 3.5% (comparée à 1319, issue du calibrage). Des résultats similaires ont été obtenus pour les autres séquences.

La figure 6 montre l'image rectifiée d'une planche de la bande dessinée, correspondant à la dernière vue de la colonne de gauche. Les angles et rapports de longueurs semblent corrects, en tenant compte du fait que la planche n'était pas parfaitement plate.

Rectification de l'image d'une case de bande dessinée (BD). Soient les six images d'une case de bande dessinée, affichées dans la figure 5 (ligne du haut), pour lesquelles nous avons estimé les matrices des homographies inter-vues. Supposons que le point principal et le facteur d'échelle de la caméra soient connus (en se ramenant éventuellement au cas $u_0 = v_0 = 0$ et $\tau = 1$). L'algorithme LIGPAL (cf. TAB. 1) peut être alors facilement adapté (seuls

FIG. 6 – À gauche : images originales ; à droite : image rectifiée après autocalibrage plan, à partir des solution exactes données en §4.3

α, β, λ et μ sont inconnus) de telle sorte qu'il soit celui de la rectification euclidienne des images de la BD (cf. §4.2). Une fois α, β, λ et μ estimés, nous appliquons la transformation Q^{-1}, donnée en (2), aux images originales (FIG. 5, ligne du haut) pour obtenir de nouvelles images rectifiées (FIG. 5, ligne du bas).

6 Conclusion

Nous avons abordé le problème de l'autocalibrage d'une caméra, à partir des seules homographies inter-vues, induites par le déplacement de la caméra observant une figure plane. La distance focale n'est pas constante et nous supposons qu'elle peut varier de façon conséquente, au cours des acquisitions. Une formulation non linéaire de ce problème, fondée sur les propriétés projectives des points cycliques a déjà été détaillée dans [10]. Pour être efficace, les solutions proposées nécessitent des valeurs initiales pour les paramètres à estimer, en particulier pour les distances focales, qui ne soient pas trop éloignées des solutions.

Notre première contribution a été de proposer une technique, simple à mettre en œuvre, qui consiste à prendre (au moins) une vue, de telle façon que le plan qui contient la figure soit (approximativement) un plan fronto-parallèle. Il est alors possible de trouver des des solutions exactes pour l'autocalibrage de la caméra, à partir de différentes hypothèses sur les paramètres intrinsèques, ont été proposées. Cette hypothèse de «fronto-parallélisme» n'est utilisée que pour initialiser les paramètres à estimer, avant d'appliquer toute méthode de résolution itérative.

La seconde contribution a été de mettre sous forme matricielle la contrainte de la ligne principale/ligne des centres, déjà utilisée dans le cadre du calibrage plan [2], qui confine le lieu du point principal à une droite. La caractéristique principale de cette contrainte est de s'exprimer indépendamment de la distance focale, ce qui peut être avantageux vis-à-vis de l'autocalibrage d'une caméra à focale variable. Nous avons montré, que lorsque le facteur d'échelle est inconnu, le lieu de la ligne principale est un faisceau de droites, de sommet le point horizon. Nous montrons que la contrainte de la ligne principale exprime aussi l'alignement de trois points, dont un est le pôle de la droite de l'infini par rapport à l'image de la conique duale aux points cycliques (ICDPC). Cet aspect est intéressant, puisque nous savons comment l'ICDPC se transforme de vue en vue, à partir des homographies inter-vues. Les expérimentations réalisées valident la performance de nos algorithmes, en termes de précision et de complexité, notamment lors de fortes variations de distance focale.

Remerciements.

Cette étude a été partiellement financée par le GdR ISIS, dans le cadre du programme «Projets jeunes chercheurs».

Références

[1] P. Gurdjos, R. Payrissat. Calibrage plan d'une caméra en mouvement à focale variable RFIA, p. 859-868, 2002.

[2] P. Gurdjos, A. Crouzil, R. Payrissat. Another Way of Looking at Plane-Based Calibration : the Centre Circle Constraint. ECCV, vol. 4, p. 252-266, 2002.

[3] P. Gurdjos, P. Sturm. Methods and Geometry for Plane-Based Self-Calibration. CVPR, vol.1, p. 491-496, 2003

[4] R. Hartley, A. Zisserman. Multiple View Geometry in Computer Vision. Cambridge University Press 2002.

[5] E. Malis, R. Cipolla. Multi-view Constraints between Collineations : Application to Self-Calibration from Unknown Planar Structures. ECCV, vol. 2, p. 610-624, 2000.

[6] Optimization Toolbox User's Guide. Version 2. The MathWorks, Inc. September 2000.

[7] C. Slama. Manual of Photogrammetry. American Society of Photogrammetry, USA, 4th edition, 1980.

[8] P. Sturm, S. Maybank. On Plane-Based Camera Calibration : a General Algorithm, Singularities, Applications. CVPR, p. 432-437, 1999.

[9] P. Sturm. Algorithms for Plane-Based Pose Estimation. CVPR, p. 1010-1017, 2000.

[10] B. Triggs. Autocalibration from Planar Scenes. ECCV, p. 89-105, 1998.

[11] Z. Zhang. A Flexible New Technique for Camera Calibration. PAMI, vol. 22, no. 11, p. 1330-1334. 2000.

[12] Z. Zhang. Camera Calibration with One-Dimensional Objects. ECCV, vol. 4, p. 161-174, 2002.