

Auto-calibrage plan d'une caméra : une interprétation simple des contraintes géométriques

Plane-based Camera Self-calibration : a Simple Interpretation of Geometric Constraints

J.F. Menudet, J.M. Becker, T. Fournel, C. Mennessier

Laboratoire Traitement du Signal et Instrumentation, UMR CNRS-UJM 5516
10 rue Barroin 42000 Saint-Etienne FRANCE
jeanfrancois.menudet@cpe.fr

Résumé

L'auto-calibrage plan vise à calculer les paramètres intrinsèques d'une caméra à partir de plusieurs images d'une même scène plane inconnue. Cet article rappelle comment ce problème a été récemment abordé par B. Triggs puis E. Malis : les deux méthodes se basent sur les homographies inter-images spécifiques à ce type de scène et conduisent à la minimisation de fonctions de coût non-linéaires. L'originalité de notre approche réside dans une interprétation géométrique simple des contraintes portant sur les homographies et la rectification métrique des images. Cette compréhension plus intuitive du problème nous permet de mieux formuler le processus d'optimisation. Les simulations effectuées montrent en effet une meilleure stabilité de notre méthode.

Mots Clef

Auto-calibrage, Caméra, Homographie, Points Cycliques, Similitude, Optimisation.

Abstract

The purpose of plane-based self-calibration is the computation of camera intrinsic parameters using a certain number of images of the same unknown planar scene. This paper recalls how this issue has been addressed by recent authors, B. Triggs and then E. Malis, using cross-views homographies specific to planar scenes, and a final optimization phase. The originality of this paper is in a straightforward geometric interpretation of constraints related to homographies and metric rectification of images. This approach provides a more intuitive formulation of the optimization process. Moreover, numerical computations show that our method achieves a better stability compared to earlier approaches.

Keywords

Self-calibration, Camera, Homography, Circular Points, Similarity, Optimization.

1 Introduction

Le calibrage d'une caméra est une tâche fondamentale dans le domaine de la vision par ordinateur dès lors qu'une interprétation métrique des images doit être obtenue. De nombreux travaux [2, 17, 19] ont peu à peu amené des solutions simples et efficaces au problème, grâce à l'utilisation de mires 3D ou planes, ces dernières apportant un avantage pratique certain (moindre encombrement, facilité de fabrication). Cependant, l'utilisation de mires représente souvent une contrainte incompatible avec les besoins applicatifs de la vision par ordinateur.

Des méthodes, dites d'auto-calibrage, sont progressivement apparues [3, 12] pour apporter la souplesse d'utilisation faisant défaut au calibrage. L'idée principale est de remplacer la mire par n'importe quelle scène 3D. La seule connaissance a priori sur la scène est celle de sa rigidité. En ajoutant éventuellement d'autres hypothèses sur la caméra, il est ainsi possible de formuler un ensemble de contraintes géométriques, à partir d'une séquence d'images. Généralement, ces algorithmes, non-linéaires par essence, souffrent d'une instabilité rendant difficile leur mise en application. En particulier, certains mouvements de caméra conduisent à une dégénérescence des contraintes [14].

Plus récemment, le concept d'auto-calibrage a été étendu aux scènes planes, omniprésentes dans l'environnement humain. L'intérêt est que la planéité de la scène représente une contrainte forte, pouvant éventuellement améliorer les résultats si elle est intégrée. La prise en compte des points cycliques du plan permet d'obtenir les équations d'auto-calibrage mises en évidence par B. Triggs [16]. Le sens de ces points et des équations associées est cependant assez peu intuitif. Une approche manipulant des objets plus simples, notamment la normale au plan, a également été introduite par E. Malis et al. [9]. Les équations qui en découlent ne semblent pas pour autant posséder une interprétation géométrique évidente.

Dans tous les cas, la résolution des équations mises en

oeuvre passe par des algorithmes d'optimisation numérique (minimisation d'une fonction de coût en l'occurrence). Or les résultats de ces algorithmes dépendent fortement du choix de la fonction à minimiser et du choix de sa paramétrisation. Une bonne compréhension de ces deux éléments est donc essentielle.

Notre contribution porte conjointement sur ces deux aspects. En premier lieu, une formulation géométrique simple de l'auto-calibrage plan nous permet de justifier la fonction de coût à minimiser. L'idée est de chercher à rectifier les images de façon à pouvoir les mettre en correspondance par des similitudes. En second lieu, une paramétrisation directe et intuitive du problème est introduite. Elle ne fait appel qu'aux paramètres intrinsèques et à la normale au plan. Tout au long de cet article, nous considéreront les paramètres intrinsèques constants afin de simplifier les notations et les explications. Cependant, toutes les notions développées se généralisent au cas d'un ou plusieurs paramètres variables (focales multiples en particulier).

Après quelques préliminaires essentiels à la compréhension du problème, nous exposerons les méthodes de B. Triggs et E. Malis. Notre contribution commencera par la présentation de l'élément central de notre approche : la rectification métrique des images. Nous mettrons alors en évidence une contrainte géométrique entre ces rectifications et les homographies inter-vues, supposées connues. Après avoir paramétré les matrices de rectifications via une certaine décomposition matricielle, nous introduirons notre formulation non-linéaire de l'auto-calibrage plan. Son lien théorique avec les approches existantes sera ensuite abordé. Enfin, nous validerons notre démarche par une comparaison expérimentale des différentes méthodes.

2 Préliminaires

2.1 Conventions

Notations : Les scalaires sont notés en italique (a, b), les vecteurs en gras ($\mathbf{n}, \mathbf{\Pi}$) et les matrices utilisent une police sans sérif (K, M). Sauf indications contraires, les vecteurs sont identifiés à des matrices colonnes. Le produit scalaire entre deux vecteurs \mathbf{a} et \mathbf{b} est donc $\mathbf{a}^T \mathbf{b}$. Le produit vectoriel de \mathbf{a} et \mathbf{b} ($\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$) est noté $\mathbf{a} \times \mathbf{b} = [\mathbf{a}]_{\times} \mathbf{b}$ avec $[\cdot]_{\times}$ la matrice anti-symétrique associée au produit vectoriel. Pour alléger la rédaction, on confondra les entités géométriques (points, plans, etc..) et les vecteurs homogènes qui les représentent (\mathbf{X} désigne le point X de coordonnées homogènes $\mathbf{X} = [X \ Y \ Z \ W]^T$). L'utilisation des coordonnées homogènes entraîne alors l'équivalence classique dite « égalité à un facteur d'échelle près », notée \sim .

Repères : L'image est munie d'un système de coordonnées (u, v) dont l'origine est le coin supérieur gauche. Les coordonnées des points 3D peuvent être exprimées soit dans un repère local \mathcal{R}_c lié à la caméra, soit dans un repère 3D global \mathcal{R}_w attaché à la scène supposée statique. Une transformation rigide (isométrie), définie par la matrice de rotation

R et la translation \mathbf{t} , permet de passer de \mathcal{R}_w à \mathcal{R}_c . Un point X de coordonnées \mathbf{X}^w dans \mathcal{R}_w a pour coordonnées $\mathbf{X}^c = \begin{bmatrix} R & \mathbf{t} \\ 0 & 1 \end{bmatrix} \mathbf{X}^w$ dans \mathcal{R}_c .

2.2 Modèle de la caméra

Nous considérons le cas du modèle sténopé pour des pixels rectangulaires. Les distorsions introduites par l'optique sont négligées. Un point de l'espace $\mathbf{X}^w = [X \ Y \ Z \ 1]^T$ se projette alors sur le plan image en un point $\mathbf{x} = [u \ v \ 1]^T$:

$$\mathbf{x} \sim K \begin{bmatrix} R & \mathbf{t} \end{bmatrix} \mathbf{X}^w \quad (1)$$

avec

$$K = \begin{bmatrix} \tau f & 0 & u_0 \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2)$$

f , τ et $\mathbf{c} = [u_0 \ v_0]^T$ sont les paramètres intrinsèques de la caméra que nous cherchons à déterminer. f est la focale (en pixels), τ le rapport d'aspect des pixels et \mathbf{c} les coordonnées dans le repère image du point principal.

On note $\bar{\mathbf{x}} = K^{-1} \mathbf{x}$ les coordonnées normalisées de \mathbf{x} qui donnent le vecteur directeur du rayon optique passant par \mathbf{x} , exprimé dans le repère de la caméra.

2.3 Géométrie de deux images d'une scène plane

Rappelons quelques propriétés essentielles propres aux images d'une figure dans un plan Π (voir figure 1).

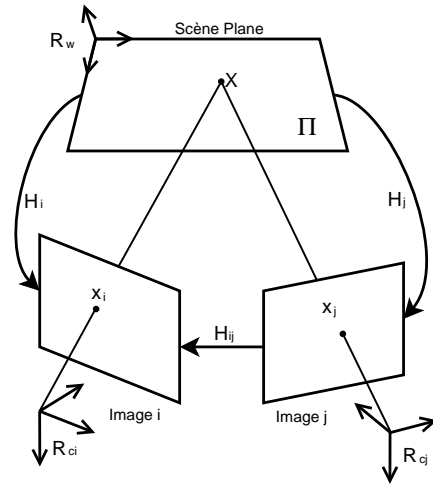


FIG. 1 – Géométrie de 2 images d'un plan : l'image i d'une scène plane peut être obtenue par l'application d'une homographie H_i . Deux images (i, j) de cette même scène sont de plus reliées par une homographie H_{ij} .

Homographie Scène-Image. On choisit le repère global de l'espace \mathcal{R}_w de façon à ce que les coordonnées d'un point X quelconque de Π soient de la forme $\mathbf{X} =$

$[X Y 0 1]^T$. On définit ainsi un repère 2D du plan dans lequel les coordonnées de ce même point sont $\tilde{\mathbf{X}} = [X Y 1]^T$. A chaque image i de la séquence est associée une matrice d'homographie 3×3 H_i permettant de transformer $\mathbf{X} \in \Pi$ en son image \mathbf{x}_i dans la vue i . Les coordonnées homogènes de \mathbf{x}_i dans le repère image sont :

$$\mathbf{x}_i \sim H_i \tilde{\mathbf{X}} \quad (3)$$

\mathcal{R}_w est lié au repère caméra par une transformation rigide (\mathbf{R}, \mathbf{t}) . En notant $(\mathbf{r}_0, \mathbf{r}_1, \mathbf{r}_2)$ les vecteurs colonnes de la matrice de rotation \mathbf{R} , on peut montrer que :

$$H_i \sim \mathbf{K} [\mathbf{r}_0 \ \mathbf{r}_1 \ \mathbf{t}] \quad (4)$$

Pour déterminer H_i , définie à un facteur d'échelle près, il suffit de connaître les coordonnées 2D de 4 points appartenant à Π et les coordonnées de leurs projections respectives dans l'image i .

Homographies inter-images. Pour chaque couple d'images (i, j) , la matrice d'homographie 3×3

$$H_{ij} = H_i H_j^{-1} \quad (5)$$

associe à un point \mathbf{x}_j de l'image j le point correspondant \mathbf{x}_i de l'image i (voir figure 1) :

$$\mathbf{x}_i \sim H_{ij} \mathbf{x}_j \quad (6)$$

Cette matrice peut être calculée à un facteur d'échelle près à partir de la correspondance d'au moins 4 points (n'ayant pas de triplet colinéaire) dans les 2 images. Il est important de noter qu'aucune information a priori sur la scène n'est nécessaire à ce calcul.

La relation (6) utilise les coordonnées des points dans le repère image (pixel). Une équation similaire existe pour les coordonnées normalisées. Il suffit de poser

$$\bar{H}_{ij} = \mathbf{K}^{-1} H_{ij} \mathbf{K} \quad (7)$$

et on obtient alors l'équivalent de (6) :

$$\bar{\mathbf{x}}_i \sim \bar{H}_{ij} \bar{\mathbf{x}}_j \quad (8)$$

2.4 Conique absolue, points cycliques et calibrage

Conique absolue. Dans l'espace projectif \mathcal{P}^3 , toutes les sphères intersectent le plan à l'infini en un unique ensemble de points. Ces points de la forme $\mathbf{X} \sim [X Y Z 0]^T = [\hat{\mathbf{X}} 0]^T$ vérifient l'équation [3] :

$$\hat{\mathbf{X}}^T \hat{\mathbf{X}} = X^2 + Y^2 + Z^2 = 0 \quad (9)$$

Mis à part le cas $X = Y = Z = 0$, cette équation n'a de solutions que si nous autorisons X, Y, Z à prendre des valeurs complexes. Cette ensemble de points à coordonnées complexes peut donc être vu comme une conique particulière du plan à l'infini, la conique absolue, associée à la matrice $\Omega \sim \mathbf{I}_3$.

Points cycliques. Les points cycliques sont l'équivalent 2D de la conique absolue. Ils sont en effet définis comme l'intersection de tous les cercles du plan et de la droite à l'infini λ_∞ d'où la contrainte $X^2 + Y^2 = 0$ sur leurs coordonnées. Les points cycliques sont par conséquent au nombre de deux et ont des coordonnées complexes conjuguées : $\mathbf{I} = [1 \ i \ 0]^T$ et $\mathbf{J} = \bar{\mathbf{I}} = [1 \ -i \ 0]^T$.

Dans le cas d'un plan Π plongé dans \mathcal{P}^3 , λ_∞ est l'intersection de Π et Π_∞ . Les points cycliques de Π sont alors les deux points d'intersection entre λ_∞ et la conique absolue (voir figure 2).

Ces points ont de plus la particularité d'être invariants par toute similitude du plan (composition d'une homothétie et d'un déplacement). Nous verrons par la suite que les similitudes planes jouent un rôle essentiel dans notre approche. Elles ont en effet la propriété de conserver les informations métriques (angles, rapport de distances) que nous recherchons dans les images.

Image de la conique absolue. Une propriété remarquable de la conique absolue est que son image ω ne dépend que des paramètres intrinsèques de la caméra :

$$\omega \sim (\mathbf{K}\mathbf{K}^T)^{-1} \sim \begin{bmatrix} 1 & 0 & -u_0 \\ 0 & \tau^2 & -\tau^2 v_0 \\ -u_0 & -\tau^2 v_0 & \tau^2 f^2 + u_0^2 + \tau^2 v_0^2 \end{bmatrix}$$

En effet, d'après (1), l'image \mathbf{x} d'un point $\mathbf{X} = [\hat{\mathbf{X}} 0]^T$ appartenant à Ω est $\mathbf{x} \sim \mathbf{K}[\mathbf{R}|\mathbf{t}]\mathbf{X} = \mathbf{K}\mathbf{R}\hat{\mathbf{X}}$. Donc :

$$\hat{\mathbf{X}} \sim \mathbf{R}^{-1}\mathbf{K}^{-1}\mathbf{x} = \mathbf{R}^T\mathbf{K}^{-1}\mathbf{x} \quad (10)$$

Ce qui donne d'après (9) :

$$\begin{aligned} \mathbf{x}^T \mathbf{K}^{-T} \mathbf{R} \mathbf{R}^T \mathbf{K}^{-1} \mathbf{x} &= 0 \\ \mathbf{x}^T (\mathbf{K}\mathbf{K}^T)^{-1} \mathbf{x} &= 0 \end{aligned} \quad (11)$$

Il est clair que que la connaissance de ω équivaut à celle de (f, τ, u_0, v_0) et donc au calibrage de la caméra [15]. De plus, si les paramètres intrinsèques sont constants, alors l'image de la conique absolue est la même dans toutes les vues (voir figure 2).

Image des points cycliques. Pour l'image d'un plan Π , une contrainte particulière existe [16] : puisque \mathbf{I} et $\mathbf{J} = \bar{\mathbf{I}}$ (points cycliques de Π) appartiennent à Ω , leurs images \mathbf{i} et \mathbf{j} doivent appartenir à ω :

$$\mathbf{i}^T \omega \mathbf{i} = \mathbf{j}^T \omega \mathbf{j} = 0 \quad (12)$$

De même que \mathbf{I} et \mathbf{J} , \mathbf{i} et \mathbf{j} sont également deux vecteurs complexes conjugués. Pour se ramener à des équations réelles, nous introduisons les vecteurs \mathbf{a} et \mathbf{b} , respectivement partie réelle et imaginaire de $\mathbf{i} = \mathbf{a} + i\mathbf{b}$. La contrainte (12) devient :

$$\mathbf{a}^T \omega \mathbf{a} = \mathbf{b}^T \omega \mathbf{b} \quad \text{et} \quad \mathbf{a}^T \omega \mathbf{b} = 0 \quad (13)$$

Ces équations sont les deux contraintes de base pour le calibrage et l'auto-calibrage d'une caméra à partir d'une scène plane.

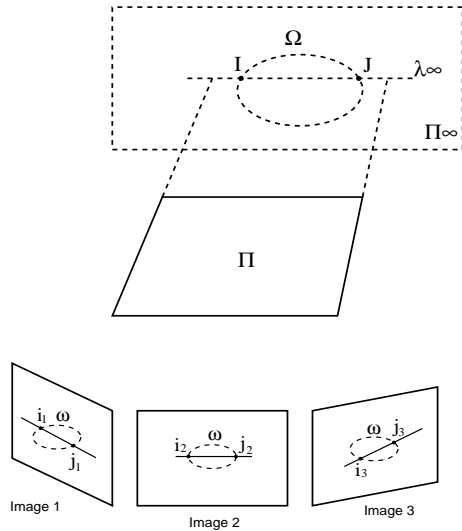


FIG. 2 – La conique absolue Ω a pour image une conique ω identique dans toutes les images. En revanche, les projections des deux points cycliques I et J du plan Π diffèrent selon les images ($i_1, j_1 ; i_2, j_2 ; \dots$)

Calibrage plan. Le problème classique du calibrage plan est le suivant : étant donnée une séquence de m images d'une *mire plane connue*, comment calculer les paramètres intrinsèques de la caméra ?

La solution décrite dans [19, 15] consiste à calculer les m homographies H_1, H_2, \dots, H_m reliant la mire et les images grâce à la *connaissance a priori de la scène*. Les images dans la vue k des points cycliques de Π sont alors connues :

$$\mathbf{i}_k \sim H_k \mathbf{I} \quad \text{et} \quad \mathbf{j}_k \sim H_k \mathbf{J} \quad (14)$$

En notant \mathbf{h}_0 et \mathbf{h}_1 les deux premiers vecteurs colonnes de H_k , il est facile de voir que les équations (13) appliquées à l'image k deviennent :

$$\mathbf{h}_0^T \omega \mathbf{h}_0 = \mathbf{h}_1^T \omega \mathbf{h}_1 \quad \text{et} \quad \mathbf{h}_0^T \omega \mathbf{h}_1 = 0 \quad (15)$$

Chaque image $k = 1..m$ donne deux contraintes linéaires sur les composantes de ω . Deux images de la même mire ($m = 2$) sont donc suffisantes pour calculer ω et calibrer la caméra (toujours sous l'hypothèse de paramètres intrinsèques constants). Si une seule image est disponible, il est malgré tout possible de retrouver les valeurs de f et τ en supposant le point principal connu. Il est également possible de prendre en compte des paramètres intrinsèques variables [15]. Dans tous les cas, les équations sont linéaires et se résolvent par le calcul d'une SVD [13, 15, 19].

La solution fournie par la SVD minimise une certaine erreur algébrique, difficilement interprétable. Dans [5], une interprétation géométrique plus intuitive est donnée aux équations du calibrage plan. Un algorithme minimisant des distances euclidiennes est alors proposé. Cette démarche permet d'améliorer significativement les résultats obtenus.

3 Auto-calibrage à partir d'une structure plane

Par rapport au cas du calibrage plan, la scène est composée d'une *structure plane inconnue*, c'est à dire d'un ensemble de n points d'intérêts coplanaires dont l'agencement est inconnu. Une séquence de m images de cette scène est acquise et nous supposons que les points d'intérêts ont été extraits et mis en correspondance dans les images.

La scène étant inconnue, il est impossible de calculer les matrices d'homographies scène-images et les équations (15) ne sont plus exploitables directement. En revanche, les correspondances de points entre les images permettent encore de calculer les matrices d'homographies inter-vues, notées H_{ij} . Ces matrices constituent les données d'entrées des méthodes d'auto-calibrage plan que nous présentons maintenant. Les paramètres intrinsèques sont toujours supposés constants.

3.1 Approche de B. Triggs [16]

La méthode présentée dans [16] s'appuie toujours sur les équations (13), faisant intervenir les projections des points cycliques dans les images. En revanche, ces projections sont désormais inconnues. L'idée principale est que si l'on connaît la projection des points cycliques dans une vue clef, il est possible de les transférer dans toutes les autres images grâce aux matrices d'homographies inter-vues. Nous supposons par la suite que l'image clef est la première de la séquence et nous noterons comme précédemment H_{k1} la matrice d'homographie de l'image 1 vers l'image k . Nous avons donc :

$$\mathbf{i}_k \sim H_{k1} \mathbf{i}_1 \quad \text{et} \quad \mathbf{j}_k \sim H_{k1} \mathbf{j}_1$$

Pour chaque image $k = 1..m$ de la séquence, (13) nous donne alors :

$$\begin{aligned} (H_{k1} \mathbf{a}_1)^T \omega (H_{k1} \mathbf{a}_1) &= (H_{k1} \mathbf{b}_1)^T \omega (H_{k1} \mathbf{b}_1) \quad (16) \\ (H_{k1} \mathbf{a}_1)^T \omega (H_{k1} \mathbf{b}_1) &= 0 \end{aligned}$$

Nous obtenons ainsi un système de $2m$ équations d'inconnues ($\omega, \mathbf{a}_1, \mathbf{b}_1$). La matrice H_{k1} peut être calculée par la mise en correspondance d'au moins 4 points dans les vues 1 et k . Les vecteurs \mathbf{a}_1 et \mathbf{b}_1 sont définis à un facteur d'échelle et une rotation 2D près [16], ce qui réduit leurs degrés de liberté de 6 à 4. La matrice $\omega = (KK^T)^{-1}$ introduit les 4 inconnues f, τ, u_0, v_0 . En théorie, 4 images ($m = 4$) suffisent donc pour déterminer l'ensemble des inconnues, 3 si le point principal est connu.

Contrairement au cas du calibrage, les équations (16) ne sont plus linéaires et il n'existe pas de méthodes directes permettant de les résoudre. L'idée est alors de reformuler le problème en tant que minimisation d'une fonction de coût, comme par exemple :

$$\min_{\omega, \mathbf{a}_1, \mathbf{b}_1} \sum_{k=1}^m \frac{(\alpha_k - \beta_k)^2 + \gamma_k^2}{\alpha_k^2 + \beta_k^2} \quad (17)$$

en posant

$$\begin{aligned}\alpha_k &= (\mathbf{H}_{k1}\mathbf{a}_1)^T\boldsymbol{\omega}(\mathbf{H}_{k1}\mathbf{a}_1) \\ \beta_k &= (\mathbf{H}_{k1}\mathbf{b}_1)^T\boldsymbol{\omega}(\mathbf{H}_{k1}\mathbf{b}_1) \\ \gamma_k &= (\mathbf{H}_{k1}\mathbf{a}_1)^T\boldsymbol{\omega}(\mathbf{H}_{k1}\mathbf{b}_1)\end{aligned}\quad (18)$$

Tous les algorithmes classiques d'optimisation numérique sont alors applicables. Le choix de l'erreur à minimiser est essentiel dans ce genre de techniques et (17) n'est pas le seul possible. Des considérations d'ordre statistique [16] peuvent par exemple conduire à la minimisation d'une fonction de coût différente. De la même façon, la paramétrisation de cette erreur en fonction des inconnues est un élément primordial.

La paramétrisation de \mathbf{a}_1 et \mathbf{b}_1 pose justement ici un problème car ces 2 vecteurs 3×1 sont définis à un facteur d'échelle et une rotation 2D près. Une solution consiste à remarquer que d'après (14), il est équivalent de rechercher $(\hat{\mathbf{i}}_1, \hat{\mathbf{j}}_1)$ ou \mathbf{H}_1 . Mais comme \mathbf{H}_1 a 8 degrés de liberté, une telle paramétrisation des points cycliques n'est pas minimale. Les 4 degrés de liberté introduits sont en fait liés à une similitude du plan. En effet, \mathbf{H}_1 peut se décomposer de la manière suivante [11, 8] :

$$\mathbf{H}_1 = \mathbf{P}_1\mathbf{S}_1 \quad (19)$$

où \mathbf{S}_1 désigne une similitude du plan et \mathbf{P}_1 décrit la partie non-métrique de \mathbf{H}_1 . Cette matrice n'a que quatre degrés de liberté et peut être mise sous la forme :

$$\mathbf{P}_1 = \begin{bmatrix} 1 & a & 0 \\ 0 & b & 0 \\ c & d & 1 \end{bmatrix} \quad (a, b, c, d) \in \mathbb{R}^4$$

Les points cycliques étant invariants par similitude, on obtient finalement :

$$\mathbf{a}_1 + i\mathbf{b}_1 \sim \mathbf{P}_1\mathbf{S}_1\mathbf{I} \sim \mathbf{P}_1\mathbf{I}$$

L'image des points cycliques est donc donnée par les deux premières colonnes de \mathbf{P}_1 . On obtient ainsi une paramétrisation minimale par le biais des 4 scalaires a, b, c, d . On remarque que la vue clef est parallèle au plan de la scène si et seulement si $a = c = d = 0$ et $b = 1$ [11]. En pratique, une telle configuration permet d'initialiser l'algorithme.

3.2 Approche d'E. Malis et al. [9]

E. Malis place au coeur de sa méthode la matrice 3×3 $\bar{\mathbf{H}}_{ij}[\mathbf{n}_j]_{\times}$ où \mathbf{n}_j est le vecteur normal au plan exprimé dans le repère de la caméra j . Il a établi que les 3 valeurs singulières de cette matrice sont $\sigma_1^{(ij)} = \sigma_2^{(ij)} > \sigma_3^{(ij)} = 0$. Sa formulation de l'auto-calibrage plan est alors :

$$\min_K \sum_{i=1}^m \sum_{j=1}^m \frac{\sigma_1^{(ij)} - \sigma_2^{(ij)}}{\sigma_1^{(ij)}} \quad (20)$$

Les vecteurs normaux au plan sont calculés par une décomposition de $\bar{\mathbf{H}}_{ij}$, similaire à celle présentée dans [1]. L'originalité de la méthode est qu'elle travaille simultanément sur toutes les homographies et ne nécessite donc pas le choix d'une vue clef. En effet, privilégier une image risque de propager au reste de la séquence l'incertitude portant sur cette vue.

4 Une nouvelle approche géométrique

4.1 Matrice de rectification

La décomposition (19) de \mathbf{H}_1 s'étend à n'importe quelle image k de la séquence, sous la forme :

$$\mathbf{H}_k = \mathbf{P}_k\mathbf{S}_k \quad (21)$$

La matrice $\mathbf{Q}_k = \mathbf{P}_k^{-1}$ permet d'obtenir une *rectification métrique* de l'image k , c'est à dire de pouvoir mesurer des angles et des rapports de longueurs sur cette image. Elle se décompose elle-même en un produit de deux matrices effectuant une rectification affine puis métrique de l'image à partir de la ligne d'horizon du plan et de la position des points cycliques sur celle-ci [8].

Les m matrices de rectifications des images de la séquence vont jouer un rôle central dans l'approche que nous proposons. Elles ont une interprétation géométrique simple et concentrent à elles seules tout l'enjeu de l'auto-calibrage comme nous allons le montrer maintenant.

4.2 Contrainte géométrique liée aux matrices de rectifications

D'après (5) et (21), $\mathbf{H}_{k1} = \mathbf{H}_k\mathbf{H}_1^{-1} = \mathbf{P}_k\mathbf{S}_k\mathbf{S}_1^{-1}\mathbf{P}_1^{-1}$.

Par suite, la matrice \mathbf{S}_{k1} définie par

$$\mathbf{S}_{k1} = \mathbf{Q}_k\mathbf{H}_{k1}\mathbf{Q}_1^{-1} = \mathbf{S}_k\mathbf{S}_1^{-1}$$

est une similitude (composition de deux similitudes).

L'interprétation géométrique de \mathbf{S}_{k1} est illustrée par la figure 3. Chaque image $k \neq 1$ rectifiée par \mathbf{Q}_k doit être *identique à une similitude près* à la rectification de l'image 1 par \mathbf{Q}_1 . La suite logique de notre démarche est de fournir un critère caractérisant « le degré de ressemblance à une similitude près » des images rectifiées. Nous pourrions alors envisager de maximiser ce critère en optimisant les m matrices de rectifications, fonctions des inconnues du problème. Pour cela, nous allons maintenant proposer une paramétrisation simple de ces matrices.

4.3 Décomposition de la matrice de rectification

Nous cherchons à établir pour \mathbf{P}_k une relation similaire à (4) ne faisant intervenir que K et \mathbf{n}_k (vecteur normal au

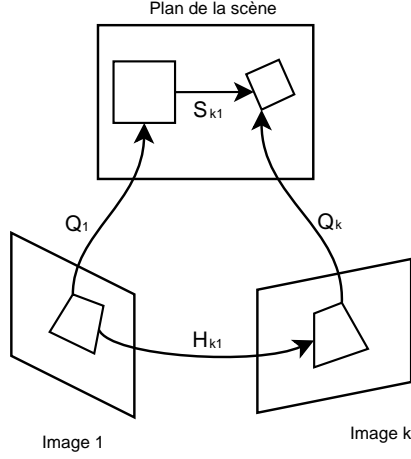


FIG. 3 – Q_1 et Q_k permettent d’obtenir respectivement une rectification métrique des images 1 et k . $S_{k1} = Q_k H_{k1} Q_1^{-1}$ est donc une similitude plane.

plan de la scène exprimé dans le repère de la caméra k). La rectification métrique $Q_k = P_k^{-1}$ de l’image k sera ainsi paramétrée simplement par ces deux éléments, à l’interprétation physique évidente.

La réduction de H_k à P_k s’effectue par le biais d’une similitude qui s’applique à la scène plane. Ceci se traduit par une réduction du nombre de degrés de liberté des paramètres extrinsèques de la caméra. La rotation et la translation dans le plan de la scène ne sont en effet plus à prendre en compte (composante euclidienne de la similitude) ainsi que la translation selon l’axe optique (facteur d’échelle de la similitude).

Plus formellement, d’après (4), la partie de H_k liée aux paramètres extrinsèques (R, t) s’écrit :

$$[\mathbf{r}_0 \ \mathbf{r}_1 \ \mathbf{t}] = R \begin{bmatrix} 1 & 0 & \vdots \\ 0 & 1 & R^T \mathbf{t} \\ 0 & 0 & \vdots \end{bmatrix} = R \begin{bmatrix} 1 & 0 & t'_x \\ 0 & 1 & t'_y \\ 0 & 0 & t'_z \end{bmatrix}$$

En utilisant les angles d’Euler, la matrice de rotation R peut se décomposer en 3 matrices de rotation, chacune dépendantes d’un seul angle : $R = ABC$. La matrice C représente en 3D une rotation autour de l’axe z . Mais appliquée à des points 2D exprimés en coordonnées homogènes, elle devient une rotation plane. En notant $\alpha = \frac{1}{t'_z}$, on obtient :

$$[\mathbf{r}_0 \ \mathbf{r}_1 \ \mathbf{t}] \sim ABC \begin{bmatrix} \alpha & 0 & \alpha t'_x \\ 0 & \alpha & \alpha t'_y \\ 0 & 0 & 1 \end{bmatrix} = R_k S_k \quad (22)$$

avec $R_k = AB$ et $S_k = C \begin{bmatrix} \alpha & 0 & \alpha t'_x \\ 0 & \alpha & \alpha t'_y \\ 0 & 0 & 1 \end{bmatrix}$.

S_k , composée d’une homothétie-translation et d’une rotation est donc une similitude plane. Quant à la matrice R_k ,

elle représente une rotation transformant le vecteur $[0 \ 0 \ 1]^T$ en \mathbf{n}_k .

Comme $H_k \sim K [\mathbf{r}_0 \ \mathbf{r}_1 \ \mathbf{t}] \sim P_k S_k$, le résultat (22) nous permet finalement d’écrire la relation recherchée

$$Q_k = P_k^{-1} \sim R_k^T K^{-1} \quad (23)$$

4.4 Obtention de contraintes algébriques

La décomposition (23) des matrices de rectifications nous permet maintenant d’obtenir une expression de S_{k1} faisant apparaître la matrice d’homographie inter-vues *normalisée* \bar{H}_{k1} , définie en (7) :

$$S_{k1} = Q_k H_{k1} Q_1^{-1} = R_k^T \bar{H}_{k1} R_1 \quad (24)$$

En tant que matrice de similitude, S_{k1} ne possède que 4 degrés de liberté au lieu des 9 habituels pour une matrice 3×3 . Nous mettons ainsi en évidence l’existence de 5 contraintes scalaires sur les inconnues $K, \mathbf{n}_1, \mathbf{n}_k$ dont dépendent \bar{H}_{k1}, R_1, R_k . Pour exprimer ces contraintes, introduisons la notation suivante pour les blocs de S_{k1} :

$$S_{k1} = \begin{bmatrix} \mathbf{r} & \mathbf{t} \\ \mathbf{z}^T & s \end{bmatrix}$$

S_{k1} est une matrice de similitude si et seulement si les contraintes suivantes sont vérifiées :

1. $s = 1$,
2. $\mathbf{z} = [0 \ 0]^T$,
3. \mathbf{r} proportionnelle à une matrice de rotation 2D.

Contrainte 1

Cette contrainte correspond au fait que \bar{H}_{k1} n’est pas définie à un facteur d’échelle près. Il est en effet montré dans [18] que si le mouvement inter-images est suffisamment général, les 3 valeurs singulières de \bar{H}_{k1} sont distinctes et la valeur médiane vaut 1. Connaissant K et H_{k1} (à un facteur près), la contrainte sur cette valeur singulière permet donc de fixer l’échelle de $\bar{H}_{k1} = K^{-1} H_{k1} K$.

Contrainte 2

Elle est en fait une conséquence de la relation suivante, mise en évidence dans [9] :

$$\mathbf{n}_k \sim \bar{H}_{k1}^{-T} \mathbf{n}_1 \quad (25)$$

En effet, par définition de R_1 et en prenant \mathbf{n}_1 unitaire, nous avons $R_1 [0 \ 0 \ 1]^T = \mathbf{n}_1$. Donc $R_1 = [\mathbf{u}_1 \ \mathbf{v}_1 \ \mathbf{n}_1]$ où \mathbf{u}_1 et \mathbf{v}_1 désignent deux vecteurs formant une base ortho-normale du plan \mathbf{n}_1^\perp . La même remarque est valable pour R_k . S_{k1} peut ainsi s’écrire

$$S_{k1} = R_k^T \bar{H}_{k1} R_1 = \begin{bmatrix} \mathbf{u}_k^T \\ \mathbf{v}_k^T \\ \mathbf{n}_k^T \end{bmatrix} \bar{H}_{k1} [\mathbf{u}_1 \ \mathbf{v}_1 \ \mathbf{n}_1]$$

Donc d'après (25) :

$$\mathbf{z} = \begin{bmatrix} \mathbf{n}_k^T \bar{\mathbf{H}}_{k1} \mathbf{u}_1 \\ \mathbf{n}_k^T \bar{\mathbf{H}}_{k1} \mathbf{v}_1 \end{bmatrix} \sim \begin{bmatrix} \mathbf{n}_1^T \mathbf{u}_1 \\ \mathbf{n}_1^T \mathbf{v}_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

car $(\mathbf{u}_1, \mathbf{v}_1, \mathbf{n}_1)$ sont les vecteurs colonnes de la matrice orthogonale \mathbf{R}_1 .

Contrainte 3

En éliminant \mathbf{n}_k grâce à (25), nous recherchons maintenant les 2 contraintes restantes sur la matrice \mathbf{S}_{k1} , fonction de $(\mathbf{H}_{k1}, \mathbf{K}, \mathbf{n}_1)$. La contrainte 2 est toujours respectée et la structure par blocs de \mathbf{S}_{k1} est désormais :

$$\mathbf{S}_{k1} = \begin{bmatrix} \mathbf{r} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}$$

Soient $\sigma_1^{(k)}$ et $\sigma_2^{(k)}$ les deux valeurs singulières de \mathbf{r} . \mathbf{S}_{k1} est une similitude si et seulement si $\sigma_1^{(k)} = \sigma_2^{(k)}$ et $\det(\mathbf{r}) > 0$.

Notons $\mathbf{D}_k = \mathbf{r}^T \mathbf{r} = \begin{bmatrix} \alpha_k & \gamma_k \\ \gamma_k & \beta_k \end{bmatrix}$ le bloc 2x2 supérieur gauche de $\mathbf{S}_{k1}^T \mathbf{S}_{k1}$ (indépendante de \mathbf{n}_k d'après l'orthogonalité de \mathbf{R}_k). Les deux valeurs propres de \mathbf{D}_k sont :

$$\lambda_{\{1,2\}}^{(k)} = \sigma_{\{1,2\}}^{(k)2} = \frac{\alpha_k + \beta_k \pm \sqrt{\Delta_k}}{2} \quad (26)$$

avec $\Delta_k = (\alpha_k - \beta_k)^2 + 4\gamma_k^2$ et $\alpha_k > 0, \beta_k > 0$.

L'égalité des valeurs singulières de \mathbf{r} est par conséquent équivalente à $\Delta_k = 0$, ce qui a lieu si et seulement si :

$$\alpha_k = \beta_k \quad \text{et} \quad \gamma_k = 0 \quad (27)$$

$\alpha_k, \beta_k, \gamma_k$ sont fonctions de \mathbf{H}_{k1} (données), \mathbf{K} et \mathbf{n}_1 (inconnues). Grâce aux deux contraintes (27) fournies par chaque image $k \neq 1$, nous pouvons finalement former un système de $2(m-1)$ équations en ces 6 inconnues que sont $f, \tau, u_0, v_0, n_x/n_z$ et n_y/n_z . Au minimum, 4 images sont donc nécessaires ($m = 4$), comme pour la méthode basée sur l'image des points cycliques.

4.5 Formulation d'une fonction de coût

En pratique, nous ne pouvons pas résoudre les $2(m-1)$ équations décrites précédemment pour deux raisons : elles ne sont pas linéaires et l'incertitude dans le calcul des \mathbf{H}_{k1} fait qu'elles ne sont jamais vérifiées exactement. Nous allons donc chercher à minimiser une fonction de coût en rapport avec ces équations.

Pour cela, revenons à l'idée de base de notre démarche : maximiser le « degré de ressemblance à une similitude près » des images 1 et k rectifiées par \mathbf{Q}_1 et \mathbf{Q}_k . Intuitivement, plus \mathbf{S}_{k1} se « rapproche » d'une similitude, plus ce degré de ressemblance est élevé. Nous proposons de caractériser la matrice \mathbf{S}_{k1} par $\eta_k = \sigma_2^{(k)}/\sigma_1^{(k)}$. Cette indice $0 \leq \eta_k \leq 1$ tend vers 1 quand \mathbf{S}_{k1} tend vers une similitude.

Pour maximiser la ressemblance des images, nous pouvons donc chercher à minimiser $(1-\eta_k^2)$, ce qui donne :

$$\min_{\mathbf{K}, \mathbf{n}_1} \sum_{k=2}^m \frac{\sigma_1^{(k)2} - \sigma_2^{(k)2}}{\sigma_1^{(k)2}} \quad (28)$$

D'après (26), nous obtenons finalement :

$$\min_{\mathbf{K}, \mathbf{n}_1} \sum_{k=2}^m \frac{\sqrt{\Delta_k}}{\alpha_k + \beta_k + \sqrt{\Delta_k}} \quad (29)$$

L'évaluation de ce critère nécessite le calcul de \mathbf{R}_1 en fonction du vecteur homogène $\mathbf{n}_1 = [n_x \ n_y \ n_z]^T$. Nous supposons qu'en pratique le cas $n_z = 0$ ne se présente pas¹ et proposons donc d'optimiser $\bar{\mathbf{n}}_1 = [\bar{n}_x \ \bar{n}_y \ 1]^T$ avec $\bar{n}_x = n_x/n_z$ et $\bar{n}_y = n_y/n_z$. En posant $\rho = 1/(1 + \|\bar{\mathbf{n}}_1\|)$, \mathbf{R}_1 peut s'exprimer en fonction de \bar{n}_x et \bar{n}_y sous la forme suivante :

$$\mathbf{R}_1 = \frac{\rho}{1-\rho} \begin{bmatrix} 1 + \rho \bar{n}_y^2 & -\rho \bar{n}_x \bar{n}_y & \bar{n}_x \\ -\rho \bar{n}_x \bar{n}_y & 1 + \rho \bar{n}_x^2 & \bar{n}_y \\ -\bar{n}_x & -\bar{n}_y & 1 \end{bmatrix}$$

4.6 Initialisation

L'initialisation des paramètres à optimiser est souvent un point délicat des algorithmes d'optimisation. Elle doit être la plus proche possible du minimum global pour éviter de converger vers un minimum local. Dans notre problème, l'initialisation peut-être solutionnée en introduisant un ou plusieurs types de connaissances a priori :

- une description euclidienne grossière de la scène.
- les valeurs typiques de certains paramètres intrinsèques.
- le parallélisme entre le plan image de la vue clef et le plan de la scène, ce qui se traduit par $\mathbf{n}_1 = [0 \ 0 \ 1]^T$ et donc $\mathbf{R}_1 = \mathbf{I}$. En choisissant le point principal au centre de l'image et $\tau = 1$, les équations (27) deviennent alors *linéaires* en f^2 (ceci est encore vrai si les focales sont différentes pour chaque image). Une approche similaire a déjà été proposée dans [6] et a prouvé son intérêt. Dans de nombreuses applications, il est en effet peu contraignant de respecter approximativement cette hypothèse de parallélisme.

4.7 Lien avec les autres méthodes

Images des points cycliques. Nous avons mis en évidence une contrainte sur \mathbf{D}_k , le bloc 2x2 supérieur gauche de $\mathbf{S}_{k1}^T \mathbf{S}_{k1}$:

$$\mathbf{D}_k = \begin{bmatrix} \alpha_k & \gamma_k \\ \gamma_k & \beta_k \end{bmatrix} \sim \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (30)$$

Or d'après les définitions de $\bar{\mathbf{H}}_{k1}$ et ω ,

$$\begin{aligned} \mathbf{S}_{k1}^T \mathbf{S}_{k1} &= \mathbf{R}_1^T \bar{\mathbf{H}}_{k1}^T \bar{\mathbf{H}}_{k1} \mathbf{R}_1 \\ &= (\mathbf{H}_{k1} \mathbf{K} \mathbf{R}_1)^T \omega (\mathbf{H}_{k1} \mathbf{K} \mathbf{R}_1) \end{aligned}$$

¹Il correspondrait à une direction de visée parallèle au plan.

Nous retrouvons donc la matrice $P_1 = KR_1$, dont les deux premières colonnes \mathbf{a}_1 et \mathbf{b}_1 donnent l'image des points cycliques dans la vue clef. En calculant $(\alpha_k, \beta_k, \gamma_k)$, les composantes de D_k , on se rend compte qu'on retrouve exactement les expressions (18). $D_k \sim I$ n'est en fait qu'une forme matricielle des équations (16)! Les contraintes algébriques que nous avons dégagées sont donc en théorie complètement équivalentes à celles déjà bien connues sur l'image des points cycliques.

Notre approche se démarque cependant sur 3 points :

- Nous donnons un sens géométrique simple et concret à la contrainte sur l'image des points cycliques : les m images rectifiées doivent fournir m reconstructions de la scène identiques à une similitude près.
- Notre paramétrage du problème ne dépend plus que des paramètres intrinsèques et de l'orientation du plan. Outre la réduction du nombre de paramètres (6 au lieu de 8) et leur sens physique évident, ceux-ci sont complètement décorrélés alors que l'image des points cycliques dépend des paramètres intrinsèques.
- La fonction de coût proposée se rapproche de la minimisation d'une erreur géométrique, même s'il est toujours difficile d'interpréter les résidus à l'issue de l'optimisation. Par rapport à (17), le critère est légèrement plus complexe à calculer (ajout d'une racine carrée) et dépend moins directement des paramètres. Le temps de calcul reste cependant comparable.

Valeurs singulières de $\bar{H}_{ij}[\mathbf{n}_j]_{\times}$. Il est impossible de ne pas évoquer les ressemblances de notre méthode avec celle de E. Malis. Premièrement, la normale au plan devient dans les 2 approches un élément central, remplaçant le concept peu intuitif de points cycliques. Ensuite, notre minimisation de l'écart entre les valeurs singulières du bloc supérieur gauche de S_{k1} est, dans l'esprit, similaire à (20). Les matrices mises en jeu sont elles-mêmes très proches dans leur forme : $R_i^T \bar{H}_{ij} R_j$ d'une part, $\bar{H}_{ij}[\mathbf{n}_j]_{\times}$ d'autre part.

Il est en fait possible de prouver que les 2 valeurs singulières du bloc supérieur gauche 2×2 de $R_i^T \bar{H}_{ij} R_j$ sont identiques à celle non-nulle de la matrice 3×3 $\bar{H}_{ij}[\mathbf{n}_j]_{\times}$. Pour cela, il faut d'abord remarquer que $[\mathbf{n}_j]_{\times} = R_j[\mathbf{k}]_{\times} R_j^T$ avec $\mathbf{k} = [0 \ 0 \ 1]^T$ (changement de base avec la matrice de passage R_j^T , changeant \mathbf{k} en \mathbf{n}_j). Or le problème initial est équivalent à prouver que $[\mathbf{n}_j]_{\times}^T \bar{H}_{ij}^T \bar{H}_{ij}[\mathbf{n}_j]_{\times}$ et le bloc supérieur gauche de $R_j^T \bar{H}_{ij}^T \bar{H}_{ij} R_j$ ont les mêmes valeurs propres non-nulles. Grâce à la première remarque et au principe de conservation des valeurs propres lors d'un changement de base $M \rightarrow P^{-1}MP$, il suffit donc de prouver que $[\mathbf{k}]_{\times}^T R_j^T \bar{H}_{ij}^T \bar{H}_{ij} R_j[\mathbf{k}]_{\times}$ et le bloc supérieur gauche de $R_j^T \bar{H}_{ij}^T \bar{H}_{ij} R_j$ ont les mêmes valeurs propres non-nulles. De part la forme de $[\mathbf{k}]_{\times}$, ce dernier point est évident.

Finalement, notre contrainte sur S_{k1} est donc également identique à celle portant sur $\bar{H}_{ij}[\mathbf{n}_j]_{\times}$. Nous fournissons par conséquent une interprétation géométrique simple de l'égalité des deux premières valeurs singulières de $\bar{H}_{ij}[\mathbf{n}_j]_{\times}$. Se-

conde conséquence, cette contrainte proposée par analogie avec la matrice essentielle [10] ne s'avère être qu'une reformulation de la méthode de Triggs basée sur l'image des points cycliques. A notre connaissance, ce lien n'avait jamais été mis en évidence.

Si en théorie toutes ces méthodes reposent sur les mêmes contraintes, il existe cependant 3 formulations différentes. Notre approche présente l'intérêt de donner une interprétation simple des contraintes algébriques mais qu'en est-il des résultats? Pour comparer les méthodes, nous proposons d'étudier une variante de [9] : si $\sigma_1^{(k1)}$ et $\sigma_2^{(k1)}$ sont les 2 premières valeurs singulières de $\bar{H}_{k1}[\mathbf{n}_1]_{\times}$, nous suggérons d'optimiser K et \mathbf{n}_1 en utilisant le critère proposé par Malis :

$$\min_{K, \mathbf{n}_1} \sum_{k=2}^m \frac{\sigma_1^{(k1)} - \sigma_2^{(k1)}}{\sigma_1^{(k1)}} \quad (31)$$

Par rapport à la méthode originale, nous introduisons une vue clef mais il n'est plus nécessaire de décomposer les matrices d'homographies pour trouver les normales au plan. En revanche, par rapport à notre formulation qui opère sur une matrice 2×2 , il faut ici calculer les valeurs singulières d'une matrice 3×3 .

5 Résultats expérimentaux

5.1 Simulation

Nous simulons une séquence de m images de n points d'intérêts coplanaires, tirés aléatoirement dans un carré de côté s (mètre). La caméra se déplace sur un arc de cercle (rayon $2m$, angle 50°) et vise le centre de ce cercle où se trouve la scène. Les m images sont régulièrement réparties sur cet arc de cercle avec la première image parallèle au plan de la scène. Pour simuler une caméra portée à la main, un mouvement aléatoire est ajouté. Les points sont ensuite projetés dans les images et perturbés par un bruit blanc gaussien d'écart type σ (en pixels). La caméra est caractérisée par ($f = 2000$ pixels, $\tau = 1$) et le point principal (u_0, v_0) est tiré aléatoirement dans un cercle centré sur le milieu de l'image et d'un diamètre de 50 pixels. Pour $s = 0.5$ m, les points occupent alors environ les 2/3 de la surface des images de taille 1024×1024 . Dans la pratique, ces valeurs sont typiques d'un appareil photo numérique grand public pour lequel les distorsions sont généralement négligeables².

Les homographies inter-vues sont calculées par SVD [19] après normalisation des données [7] et l'initialisation de f est obtenue en utilisant l'hypothèse de parallélisme de la vue clef (cf. 4.6). Du fait du mouvement aléatoire de la caméra, cette hypothèse n'est cependant pas exactement respectée (l'angle entre la normale au plan et l'axe optique peut atteindre 15°). La minimisation des différentes fonctions de coût est effectuée par une méthode de quasi-Newton type BFGS [4].

²Au moins pour les 2/3 supérieurs de la plage focale du zoom.

Les tests présentés se concentrent sur le calcul de f en faisant l’hypothèse $\tau = 1$ et $u_0 = v_0 = 512$. Dans notre simulation, la première hypothèse est donc vérifiée ce qui n’est pas le cas de la seconde. Nous testons les 3 méthodes suivantes :

- ALG_PC : minimisation de (17) en optimisant $f, \mathbf{a}_1, \mathbf{b}_1$
- SVD_N : minimisation de (31) en optimisant f, \mathbf{n}_1
- SIM_N : minimisation de (29) en optimisant f, \mathbf{n}_1

La figure (4) montre l’influence de la précision avec laquelle sont détectés 25 points d’intérêts contenus dans un carré de 50 cm de côté. Les résultats présentés sont la moyenne de 1000 tirages pour lesquels l’erreur relative sur f est calculée. L’hypothèse de parallélisme de la vue clef permet d’obtenir une initialisation de la focale avec une erreur inférieure à 10%. Pour $\sigma = 0$, l’erreur sur f n’est pas nulle à cause de l’hypothèse non vérifiée sur la position du point principal. Elle reste cependant faible. L’erreur augmente ensuite linéairement pour les 3 algorithmes, mais elle est constamment plus faible (de l’ordre de 20%) pour notre méthode (SIM_N) et la variante de Malis (SVD_N).

La convergence est obtenue après une quarantaine d’itérations pour ALG_PC et une trentaine pour SIM_N et SVD_N. La réduction du nombre de paramètres explique ce phénomène bien que la fonction de coût soit vraisemblablement « moins convexe ». Nous pensons que cette bonne adéquation entre la fonction de coût et son paramétrage explique les meilleures performances de SIM_N et SVD_N. Ces deux algorithmes donnent par ailleurs des résultats quasiment identiques ce qui n’est pas étonnant puisque leur seule différence réside dans une légère modification de la fonction de coût.

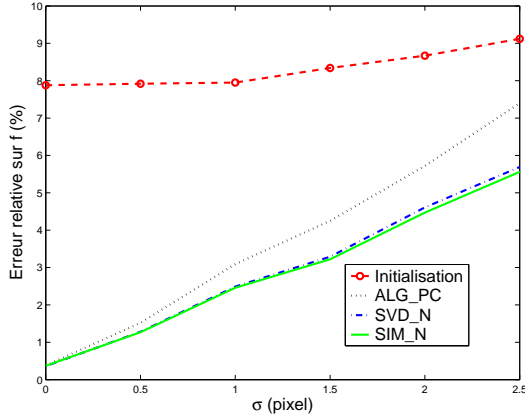


FIG. 4 – Résultats des 3 méthodes selon le bruit de mesure.

5.2 Données réelles

Une séquence de 5 images (résolution 1600x1200) est acquise par un appareil photo numérique récent dont l’auto-focus a été débrayé afin de garder les paramètres intrinsèques constants. La scène est constituée d’une mire plane connue type damier (voir figure 5) dont les 64 coins sont

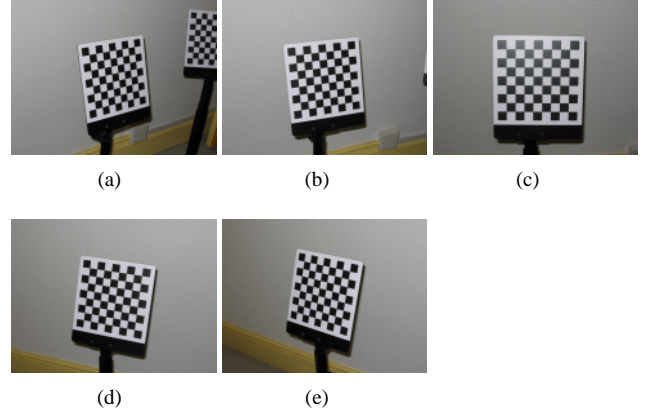


Figure 5: Séquence de 5 images d’une figure plane.

mis en correspondance dans les images avec une précision de l’ordre de 0.15 pixels. Les homographies entre l’image (c), supposée parallèle au plan, et les images (a,b,d,e) sont ensuite calculées comme précédemment. Pour chacune des 3 méthodes, le tableau 1 montre les résultats obtenus en recherchant la focale seule ou la focale plus le point principal. La connaissance a priori sur la scène permet également de calculer les homographies scène-image et donc de calibrer l’appareil selon la méthode décrite dans [15]. CAL désigne les valeurs obtenues de cette façon et serviront de références. Les paramètres non-optimisés sont entre parenthèses et prennent leur valeur par défaut. Dans tous les cas, l’hypothèse $\tau = 1$ est utilisée. Le parallélisme de la vue (c) permet d’initialiser la focale avec une erreur d’environ 5% et le point principal est placé au milieu de l’image.

TAB. 1 – Paramètres intrinsèques estimés par les différentes méthodes. Les valeurs entre parenthèses sont les valeurs par défaut des paramètres non-optimisés.

	f	τ	u_0	v_0
CAL	2934.5	0.998	774.0	601.4
ALG_PC	2928.6 2945.9	(1) (1)	(800) 827.9	(600) 548.2
SVD_N	2933.7 2929.3	(1) (1)	(800) 866.5	(600) 577.9
SIM_N	2933.8 2929.8	(1) (1)	(800) 865.5	(600) 576.6

Dans tous les cas, la focale estimée est très proche de celle fournie par le calibrage ($< 0.4\%$) en particulier lorsque le point principal n’est pas recherché ($< 0.2\%$). Lorsqu’il est mis en jeu dans l’optimisation, sa position après convergence se révèle être assez éloignée de celle donnée par le calibrage. Le faible nombre d’images (8 contraintes pour 5 ou 7 inconnues selon la méthode) explique cette imprécision mais nous montrons qu’il est possible d’avoir une bonne estimation de la focale avec seulement 5 images.

Afin de valider ces résultats d'une autre façon, nous avons également mesuré sur les images rectifiées les angles du quadrilatère englobant la mire. Quelque soit la méthode, les angles mesurés sont bien égaux à $90^\circ \pm 0.1^\circ$. Nous confirmons ainsi l'intérêt des algorithmes d'auto-calibrage plan qui permettent d'obtenir de façon très simple des informations métriques fiables.

6 Conclusion et perspectives

Notre approche de l'auto-calibrage plan repose sur la connaissance des matrices d'homographies inter-vues. Notre principale contribution est de donner une interprétation géométrique simple du problème, en termes de rectifications métriques des images. Nous montrons ainsi que les méthodes proposées par B. Triggs [16] et E. Malis [9] reposent en fait sur les mêmes contraintes et diffèrent seulement par la formulation du problème d'optimisation. L'algorithme que nous présentons donne expérimentalement des résultats comparables à une variante de [9] tout en étant plus simple et surtout montre une stabilité accrue par rapport à [16].

Plusieurs aspects méritent d'être développés. En premier lieu, l'un de nos objectifs est de parvenir à la minimisation d'un critère ayant un sens géométrique concret (par exemple la conservation des angles dans les images rectifiées). Ensuite, notre méthode permet facilement de prendre en compte des paramètres intrinsèques variables. Les prochains tests à effectuer concernent donc ce cas de figure. Enfin, il serait intéressant d'étudier dans quelle mesure la « contrainte de la ligne principale » [6] peut s'intégrer dans notre démarche. Elle permet en effet d'obtenir une formulation non-linéaire indépendante de la focale, conduisant à un problème d'optimisation a priori plus simple.

Références

- [1] O. Faugeras et F. Lustman, Motion and structure from motion in a piecewise planar environment, *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 2, pp. 485-508, 1988.
- [2] O. Faugeras et G. Toscani, The Calibration problem for stereo, In *Proc. IEEE Int. Conf. on Computer Vision and Pattern Recognition*, pp. 15-10, 1999.
- [3] O. Faugeras, Q.-T. Luong et S. J. Maybank, Camera self-calibration : Theory and Experiments, In *European Conference on Computer Vision*, Santa Margherita Ligure, Italie, Mai 1992. Springer-Verlag.
- [4] P. E. Gill, W. Murray et M. H. Wright, Practical Optimization, *Academic Press*, Londres, 1989.
- [5] P. Gurdjos, A. Crouzil et R. Payrissat, Another Way of Looking at Plane-Based Calibration : the Centre Circle Constraint, In *Proc. of the 7th European Conference on Computer Vision*, vol. 4, pp. 252-266, 2002
- [6] P. Gurdjos et P. F. Sturm, Methods and Geometry for Plane-Based Self-Calibration, In *Proc. of the International Conference on Computer Vision and Pattern Recognition*, Madison, Wisconsin, USA, vol. 1, pp. 491-946, juin 2003.
- [7] R. I. Hartley, In Defense of the Eight-Point Algorithm, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, pp. 580-593, juin 1997.
- [8] D. Liebowitz, Camera Calibration and Reconstruction of Geometry from Images, *Thèse de l'université d'Oxford*, 2001.
- [9] E. Malis et R. Cipolla, Camera self-calibration from unknown planar structures enforcing the multi-view constraints between collineations, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, pp. 1268-1272, Septembre 2002.
- [10] P. Mendonca and R. Cipolla, A simple technique for self-calibration, In *IEEE Int. Conf. on Comp. Vision and Pattern Recognition*, pp. 500-505, 1999.
- [11] E. Noirfalise, F. Jurie et P. Sturm, Autocalibration itérative de caméras à partir de structures planes, *Congrès de Reconnaissance des Formes et Intelligence Artificielle*, Toulouse, France, Janvier 2004.
- [12] M. Pollefeys et L. Van Gool, A stratified approach to metric self-calibration, In *IEEE Conf. Computer Vision & Pattern Recognition*, San Juan, Porto Rico, pp. 407-412, Juin 1997.
- [13] D. Serre, Les Matrices, *Dunod*, 2001.
- [14] P. F. Sturm, Critical motion sequences for monocular self-calibration and uncalibrated Euclidean reconstruction, *IEEE International Conference on Computer Vision and Pattern Recognition*, San Juan, Porto Rico, pp. 1100-1105, Juin 1997.
- [15] P. F. Sturm et S. J. Maybank, On Plane-Based Camera Calibration : A General Algorithm, Singularities, Applications, *CVPR-IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 432-437, Fort Collins, Colorado, USA, Juin 1999.
- [16] B. Triggs, Autocalibration from planar sequences, In *Proceedings of 5th European Conference on Computer Vision*, Freiburg, Allemagne, Juin 1998.
- [17] R. Y. Tsai, A Versatile Camera Calibration Technique For High Accuracy 3-D Machine Vision Metrology Using Off-The-Shelf TV Cameras And Lenses, *IEEE Journal Of Robotics And Automation*, vol. Ra-3, pp. 323-344, août 1987.
- [18] Z. Zhang et A.R. Hanson, Scaled euclidean 3D reconstruction based on externally uncalibrated cameras, *IEEE Symposium on Computer Vision*, pp. 37-42, 1995.
- [19] Z. Zhang, A flexible new technique for camera calibration, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, pp. 1330-1334, 2000.