

Calibrage de caméra et reconstruction 3D à l'aide de parallélépipèdes

Marta Wilczkowiak* Edmond Boyer Peter Sturm

MOVI–GRAVIR–INRIA Rhône-Alpes, 655 Avenue de l'Europe, 38330 Montbonnot

Marta.Wilczkowiak@inrialpes.fr, Edmond.Boyer@inrialpes.fr, Peter.Sturm@inrialpes.fr

Résumé

Dans cet article, nous nous intéressons à l'utilisation des parallélépipèdes pour le calibrage de caméras et la reconstruction de scènes tridimensionnelles. Les parallélépipèdes permettent d'exploiter, de manière naturelle, les contraintes géométriques fréquemment présentes dans une scène : le parallélisme et l'orthogonalité par exemple. Un sous ensemble des parallélépipèdes – les cuboïdes¹ – ont déjà été utilisés par le passé pour le calibrage partiel de caméras, mais le potentiel qu'offrent les parallélépipèdes en perception tridimensionnelle n'a jamais été clairement établi. Dans ce document, nous mettons en évidence ce potentiel au travers d'une étude approfondie des contraintes fournies par les parallélépipèdes. Nous montrons en particulier la dualité qui existe entre les caractéristiques intrinsèques d'une caméra et celles d'un parallélépipède. Pour illustrer cette étude, nous présentons par ailleurs une application interactive permettant la reconstruction d'une scène à partir d'une seule image. Dans cette application, seul un faible nombre de connaissances sur la scène sont nécessaires, et un modèle tridimensionnel peut rapidement être obtenu à partir d'une seule image non calibrée.

Mots-clés : *calibrage, reconstruction, parallélépipède, mono-image.*

1 Introduction

Dans cet article, nous décrivons comment exploiter certaines contraintes géométriques, à l'aide de parallélépipèdes, pour calibrer des images et en extraire des informations tridimensionnelles sur la scène. Nous proposons pour cela une approche simple et pratique, qui ne nécessite que peu d'interactions et de connaissances *a priori*. Le but est d'obtenir une application qui permette de modéliser des scènes tridimensionnelles à partir de photos prises avec des caméras quelconques, comme c'est le cas par exemple avec des images prises sur le réseau Internet ou avec des photos amateurs.

Plusieurs approches existantes utilisent des contraintes géométriques naturelles pour calibrer une caméra. Caprile et Torre [2] ont les premiers proposé une méthode d'estimation de certains paramètres intrinsèques de la caméra basée sur des cuboïdes. Les points de fuite définis par les projections des arêtes d'un cuboïde dans l'image sont utilisés pour cela. Ces points de fuite correspondent à trois directions orthogonales dans l'espace tridimensionnel, ce qui fournit des contraintes sur la transformation entre la scène et sa projection dans l'image. Les points de fuite ont aussi été utilisés par de nombreux chercheurs pour le calibrage partiel de caméras et la reconstruction de scènes [9, 5, 11, 6, 8]. Malheureusement le calcul des points de fuite dans une image est une opération difficile car numériquement instable [11]. De plus, le traitement indépendant de chaque point de fuite ne permet pas d'utiliser pleinement l'information contenue dans une structure parallélépipédique. Dans cet article, nous montrons comment exploiter cette information, au travers d'une matrice de projection, sans calcul explicite des points de fuite.

*Ce travail a été soutenu par le projet IST-1999-10756, VISIRE.

1. Le terme cuboïde dans ce document désigne un parallélépipède rectangle.

Récemment, certaines approches utilisant des structures planaires et leurs homographies ont été proposées [15, 18]. Bien que plus flexibles que les méthodes standard de calibrage, ces approches nécessitent soit des informations sur les positions dans l'espace des points de référence, soit plusieurs images en positions générales. De plus, les structures planaires utilisées pour le calibrage constituent souvent les faces d'un parallélépipède, qui n'est pas complètement exploité. Notre approche peut donc être considérée comme une généralisation des méthodes à base d'homographies aux structures parallélépipédiques. Les informations euclidiennes qui sont nécessaires sont moins contraignantes que celles nécessaires aux méthodes à base d'homographies. Nous verrons, en particulier, que deux angles droits et quelques hypothèses sur la caméra sont suffisants pour calibrer de manière satisfaisante.

Citons enfin quelques travaux qui utilisent directement les parallélépipèdes. Dans [7], des cuboïdes sont utilisés pour reconstruire une scène, mais plusieurs vues sont nécessaires. Dans [4], les parallélépipèdes sont utilisés dans une application de réalité augmentée. La méthode proposée présente l'inconvénient d'être basée sur la résolution d'équations non linéaires et d'imposer un modèle de caméra clairement limité (rapport d'échelle unitaire).

Notre approche est basée sur la dualité entre les caractéristiques intrinsèques de la caméra et d'un parallélépipède. Dans une projection image d'un parallélépipède, la connaissance d'un paramètre de la caméra ou du parallélépipède fournit une contrainte sur les paramètres de l'ensemble de paramètres dual (parallélépipède ou caméra). Les parallélépipèdes constituent donc un outil flexible qui s'adapte facilement à de nombreuses scènes d'environnements humains. Les structures parallélépipédiques sont, en effet, souvent présentes dans les constructions humaines (toits, façades, etc.), et une variété importante d'images peuvent être traitées facilement à partir de cette approche. De plus, en ajoutant aux parallélépipèdes des contraintes géométriques sur les points de la scène telles que la coplanarité, la colinéarité et l'appartenance à un parallélogramme, il est possible d'obtenir très facilement un modèle réaliste de la scène. Dans ce document, nous présentons un système interactif de reconstruction à partir d'une seule image basé sur ces principes.

2 Définitions

Modèle de caméra. Nous utilisons le modèle de caméra *sténopé* (projection perspective). Dans ce modèle, la projection p dans le plan image du point P de l'espace tridimensionnel est décrite par la relation projective : $p \sim M \cdot P$, où \sim représente la relation d'égalité à un facteur d'échelle près, M est une matrice de projection 3×4 qui peut être décomposée en : $M = K \cdot [R \ t]$, où $[R \ t]$ est la matrice 3×4 déterminant l'orientation relative R^T et la position $-R^T t$ de la caméra dans l'espace et K est la matrice des paramètres intrinsèques de la caméra :

$$K = \begin{bmatrix} \alpha_u & s = 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (1)$$

où α_u et α_v représentent la distance focale de la caméra exprimée suivant les tailles horizontale et verticale du pixel, s le facteur d'obliquité, considéré ici comme nul (mais cette contrainte peut être relâchée) et (u_0, v_0) les coordonnées en pixels de l'intersection de l'axe optique avec le plan image (point principal).

Les parallélépipèdes. Un parallélépipède est défini par douze paramètres : six paramètres externes décrivant son orientation et sa position dans l'espace, et six paramètres internes décrivant ses caractéristiques euclidiennes – trois longueurs d'arêtes et trois angles entre les directions des arêtes (voir la figure 1-(a)). La projection perspective d'un parallélépipède est, à un facteur d'échelle près, définie par onze paramètres. En conséquence, cinq points et une direction du plan image sont suffisants pour caractériser l'image d'un parallélépipède. Les projections des autres sommets du parallélépipède peuvent alors être déterminées par une simple règle perspective à trois points de fuite, en utilisant le parallélisme des arêtes du parallélépipède

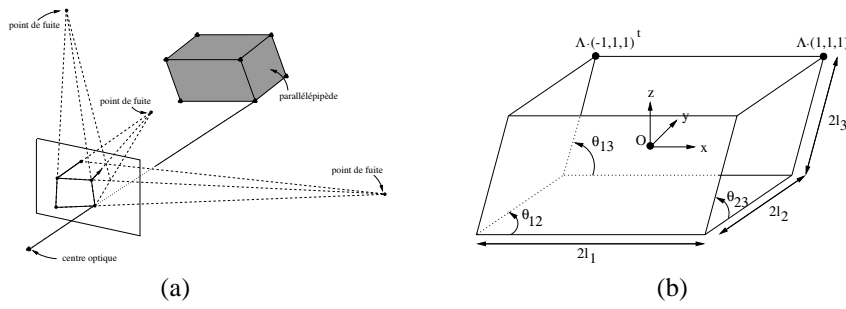


FIG. 1- (a) Un parallélépipède et sa projection définie uniquement par cinq points et une direction dans l'image. Les projections des arêtes du parallélépipède définissent trois points de fuite dans l'image. (b) La paramétrisation du parallélépipède : $2l_i$ (les longueurs des arêtes); θ_{ij} (les angles entre les arêtes du parallélépipède).

(voir la figure 1-(b)). Il en résulte que le parallélisme ne fournit aucune contrainte sur la projection et n'est donc pas utile pour le calibrage. Ce dernier est basé sur la connaissance d'informations euclidiennes sur le parallélépipède : les rapports de longueurs et les angles entre arêtes. Les projections additionnelles de sommets (en plus des cinq et demie citées) sont utiles pour stabiliser numériquement les résultats, mais ne fournissent pas de contraintes nouvelles pour le calibrage.

3 Calibrage à partir de parallélépipèdes

Les méthodes classiques utilisant les contraintes de parallélisme et d'orthogonalité consistent habituellement à déterminer les points de fuite, puis à exploiter le fait que les directions correspondantes dans l'espace sont orthogonales [2, 8]. Les points de fuite sont généralement estimés en cherchant des intersections de droites dans l'image, ou, plus précisément, les points qui minimisent la distance aux projections des droites parallèles [2, 10, 11]. De telles méthodes génèrent une forte instabilité numérique, en particulier lorsque les arêtes projetées dans l'image sont quasi-parallèles.

Contrairement à ces méthodes, nous ne calculons pas directement les points de fuite, mais nous estimons la meilleure matrice de projection telle que les points définis dans l'image correspondent aux projections des sommets d'un parallélépipède. Cette matrice nous fournit des contraintes sur la matrice des paramètres intrinsèques de la caméra. La forme de ces contraintes dépend de l'information disponible sur les paramètres du parallélépipède, mais elles sont, en général, linéaires. À ces équations peuvent, bien entendu, s'ajouter d'autres équations linéaires provenant de connaissances *a priori* sur la caméra.

3.1 Matrice de projection d'un parallélépipède

Soit un parallélépipède dont les paramètres sont $R, R^T t, (l_1, l_2, l_3)$ et $(\theta_{12}, \theta_{23}, \theta_{13})$ – les paramètres définissant, respectivement, son orientation, sa position, ses longueurs et angles. On peut associer à ce parallélépipède un repère affine, dans lequel les coordonnées des sommets sont de la forme $\{-1, 1\}^3$ (Fig. 1-(b)). La relation projective de passage du repère affine associé au parallélépipède vers le repère euclidien correspondant s'exprime alors sous la forme :

$$\tilde{\Lambda} = \begin{pmatrix} l_1 & l_2 c_{12} & l_3 c_{13} & 0 \\ 0 & l_2 s_{12} & l_3 \frac{c_{23} - c_{13} c_{12}}{s_{12}} & 0 \\ 0 & 0 & l_3 \sqrt{\frac{s_{12}^2 - c_{13}^2 s_{12}^2 - (c_{23} - c_{13} c_{12})^2}{s_{12}^2}} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (2)$$

où : $c_{ij} = \cos \theta_{ij}$, $s_{ij} = \sin \theta_{ij}$, $\theta_{ij} \in]0 \pi[$, $l_i > 0$. À noter que dans le cas d'un cuboïde la matrice $\tilde{\Lambda}$ est alors diagonale.

Les projections des sommets d'un parallélépipède dans l'image $p_{i \in [1..8]} \sim [\alpha_i u_i, \alpha_i v_i, \alpha_i]^T$ doivent satisfaire l'expression suivante :

$$\begin{pmatrix} \alpha_1 u_1 & \alpha_2 u_2 & \alpha_3 u_3 & \dots & \alpha_8 u_8 \\ \alpha_1 v_1 & \alpha_2 v_2 & \alpha_3 v_3 & \dots & \alpha_8 v_8 \\ \alpha_1 & \alpha_2 & \alpha_3 & \dots & \alpha_8 \end{pmatrix} = \tilde{X} \cdot \begin{pmatrix} 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}, \quad (3)$$

où \tilde{X} est la matrice de projection 3×4 définie à un facteur d'échelle γ près par :

$$\tilde{X} = \gamma M \cdot \tilde{\Lambda} = \gamma K \cdot [R \ t] \cdot \tilde{\Lambda}. \quad (4)$$

La matrice \tilde{X} est la composée d'une matrice de projection M et d'une matrice de transformation affine $\tilde{\Lambda}$. Il n'existe pas, en conséquence, de contrainte sur la forme de \tilde{X} telle que l'orthonormalité. Elle peut être estimée directement à partir des contraintes fournies par l'équation (3). Six projections p_i de sommets d'un parallélépipède (ou cinq points et une direction) sont alors suffisantes pour résoudre, à un facteur d'échelle près, le système linéaire d'équations (nous utilisons pour cela une méthode basée sur une décomposition en valeurs singulières [14]).

3.2 Dualité caméra-parallélépipède

La matrice de projection \tilde{X} , définie dans le paragraphe précédent, contient toute l'information géométrique fournie par la projection du parallélépipède.

Soit $X = \gamma K \cdot R \cdot \Lambda$ et Λ les matrices contenant les trois premières lignes et colonnes des matrices \tilde{X} et $\tilde{\Lambda}$. Soit $\omega = K^{-T} \cdot K^{-1}$ la matrice de la projection image de la conique absolue, cette matrice caractérisant les propriétés intrinsèques de la caméra. Et enfin, soit $\mu = \Lambda^T \cdot \Lambda$ l'entité analogue qui caractérise les propriétés intrinsèques du parallélépipède. Les matrices ω et μ sont de la forme :

$$\omega = \begin{pmatrix} \frac{1}{\alpha_u^2} & 0 & -\frac{u_0}{\alpha_u^2} \\ 0 & \frac{1}{\alpha_v^2} & -\frac{v_0}{\alpha_v^2} \\ -\frac{u_0}{\alpha_u^2} & -\frac{v_0}{\alpha_v^2} & (1 + \frac{u_0^2}{\alpha_u^2} + \frac{v_0^2}{\alpha_v^2}) \end{pmatrix}, \mu = \begin{pmatrix} l_1^2 & l_1 l_2 \cos \theta_{12} & l_1 l_3 \cos \theta_{13} \\ l_1 l_2 \cos \theta_{12} & l_2^2 & l_2 l_3 \cos \theta_{23} \\ l_1 l_3 \cos \theta_{13} & l_2 l_3 \cos \theta_{23} & l_3^2 \end{pmatrix} \quad (5)$$

À partir de ces expressions et de l'équation (4), on obtient l'expression suivante :

$$X^T \cdot \omega \cdot X = \gamma^2 \mu. \quad (6)$$

L'expression ci-dessus illustre la dualité existante entre les caractéristiques intrinsèques d'une caméra et celles d'un parallélépipède. En effet, la matrice X ne dépend que des projections des sommets du parallélépipède dans l'image. Une fois ces projections définies, les caractéristiques de la caméra et celles du parallélépipède sont mutuellement dépendantes. Fixer une caractéristique de l'un des ensembles engendre une contrainte sur l'autre ensemble de paramètres. Notons X_i la i -ème colonne de la matrice X . Les contraintes typiques utilisées pour le calibrage sont alors :

1. un angle connu θ_{ij} donne une contrainte quadratique : $(X_i^T \cdot \omega \cdot X_j)^2 - \cos^2 \theta_{ij} (X_i^T \cdot \omega \cdot X_i) (X_j^T \cdot \omega \cdot X_j) = 0$;
2. un angle *droit* connu θ_{ij} donne une contrainte linéaire : $X_i^T \cdot \omega \cdot X_j = 0$;
3. un rapport de longueurs connu $\frac{l_i^2}{l_j^2}$ donne une contrainte linéaire : $X_i^T \cdot \omega \cdot X_i - \frac{l_i^2}{l_j^2} X_j^T \cdot \omega \cdot X_j = 0$.

À noter :

- Les colonnes X_i de X représentent en effet les coordonnées homogènes des points de fuite associés aux directions du parallélépipède.

- L'ensemble des contraintes existantes sur ω (les équations dérivées à partir de plusieurs primitives ou celles fournies par la connaissance des paramètres intrinsèques de la caméra) peuvent être traitées de manière simultanée. Cela est particulièrement intéressant dans le cas linéaire.
- La connaissance d'une longueur d'arête l_i ne donne pas de contrainte sur les paramètres intrinsèques de la caméra. Par contre, une telle information permet de s'affranchir du facteur d'échelle pour la reconstruction, et d'obtenir des positions exactes pour les points de la scène.

Dans la suite de cet article, nous traitons les contraintes linéaires qui permettent un calcul direct des paramètres intrinsèques. Les contraintes quadratiques, un angle connu non droit par exemple, peuvent être ensuite ajoutées pour stabiliser la résolution effectuée, ou pour trouver un nombre fini de solutions dans le cas où les contraintes linéaires sont trop peu nombreuses.

3.3 Estimation des paramètres intrinsèques de la caméra

Supposons que l'on dispose d'au moins quatre contraintes linéaires telles quelles sont décrites dans le paragraphe précédent. Ces contraintes provenant d'une ou plusieurs parallélépipèdes, il est alors possible d'estimer les éléments de la matrice ω à un facteur d'échelle près, par une méthode des moindres carrés linéaire classique [14]. Les paramètres intrinsèques de la caméra sont ensuite déterminés à partir de ω de la manière suivante :

$$u_0 = -\frac{\omega_{13}}{\omega_{11}}, \quad v_0 = -\frac{\omega_{23}}{\omega_{22}}, \quad \alpha_u^2 = \frac{\omega_{11}\omega_{22}\omega_{33} - \omega_{11}\omega_{23}^2 - \omega_{22}\omega_{13}^2}{\omega_{11}^2\omega_{22}}, \quad \alpha_v^2 = \frac{\omega_{11}}{\omega_{22}}\alpha_u^2. \quad (7)$$

Connaissances et hypothèses sur la caméra. Certaines informations sur la caméra peuvent soit être disponibles : un rapport d'échelle α_u/α_v fourni par le fabricant de la caméra par exemple, soit découler d'hypothèses faites sur la caméra : la position du point principal au centre de l'image par exemple, cette dernière hypothèse ayant notamment été expérimentée avec succès dans notre approche (voir le paragraphe 3.5). Dans ce paragraphe, nous montrons comment inclure ces informations dans le système de calibrage. Cela permet d'éviter le calcul de paramètres déjà connus, ainsi que de calibrer la caméra dans les cas où l'on ne dispose pas d'un nombre suffisant de contraintes.

La matrice K des paramètres intrinsèques de la caméra peut être décomposée en deux matrices : $K = K_k \cdot K_u$ caractérisant respectivement les paramètres connus K_k et inconnus K_u . Les coordonnées des points dans l'image p_i peuvent alors être transformées en utilisant la matrice des paramètres connus : $p'_i = K_k^{-1} \cdot p_i$. La matrice de projection \tilde{X}' dépendant des p'_i s'exprime ensuite de la manière suivante :

$$\tilde{X}' = K_k^{-1} \cdot \gamma K \cdot [R \ t] \cdot \tilde{\Lambda} = \gamma K_u \cdot [R \ t] \cdot \tilde{\Lambda}. \quad (8)$$

Cette matrice peut être estimée à partir de points p'_i en utilisant une approche similaire à celle décrite pour \tilde{X} . Cela conduit à des expressions simplifiées pour les éléments de la matrice $\omega = K_u^{-T} \cdot K_u^{-1}$. Par exemple, si le rapport d'échelle $\tau = \alpha_u/\alpha_v$ est connu, alors :

$$K_k = \begin{pmatrix} \tau & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad K_u = \begin{pmatrix} \alpha_v & 0 & \frac{u_0}{\tau} \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \omega = \begin{pmatrix} \frac{1}{\alpha_v^2} & 0 & -\frac{u_0}{\tau\alpha_v^2} \\ 0 & \frac{1}{\alpha_v^2} & -\frac{v_0}{\alpha_v^2} \\ -\frac{u_0}{\tau\alpha_u^2} & -\frac{v_0}{\alpha_v^2} & \left(1 + \frac{u_0^2}{\tau^2\alpha_u^2} + \frac{v_0^2}{\alpha_v^2}\right) \end{pmatrix}.$$

De cette façon, chaque information sur les paramètres intrinsèques de la caméra peut être prise en compte. Dans le cas où ces informations sont peu fiables (hypothèses approximatives), il est alors possible

d'utiliser une approche alternative qui consiste à ajouter des équations supplémentaires au système de calibrage (e.g.: $\tau = \tau_0$), puis à résoudre l'ensemble, toujours suivant une méthode des moindres carrés. Cette solution offre l'avantage d'ajuster au mieux l'ensemble des paramètres et de ne pas imposer de valeurs exactes. En fonction de la "qualité" des informations disponibles sur la caméra, l'une, l'autre ou bien un mélange des deux solutions seront utilisées.

3.4 Singularités

La plupart des algorithmes de calibrage et d'auto-calibrage présentent des configurations dégénérées pour lesquelles la détermination des paramètres des caméras est impossible. Une conséquence directe est que, dans les situations proches de ces cas singuliers, les résultats du calibrage seront numériquement instables. Dans ce paragraphe, nous examinons en détail les singularités propres à l'algorithme linéaire introduit précédemment (voir le paragraphe 3.2). Nous présentons tout d'abord les principes de l'étude effectuée, puis les singularités existantes dans les principaux cas pratiques : les parallélépipèdes avec deux et trois angles droits.

Une singularité, dans le contexte qui est le nôtre, signifie que l'équation (6) possède plusieurs solutions pour les matrices ω et μ . L'existence d'une telle singularité dépend de l'orientation relative de la caméra et du parallélépipède, ainsi que des connaissances sur leurs paramètres intrinsèques. Par contre, elle ne dépend pas de la position relative de la caméra et du parallélépipède. Soit $K = K_k \cdot K_u$ la vraie matrice des paramètres intrinsèques de la caméra et $K' = K_k \cdot K'_u$ la matrice estimée (les matrices K et K' , décomposées comme dans le paragraphe 3.3 en une partie connue et une inconnue, ont évidemment la même partie connue). De même, pour le parallélépipède, $\Lambda = \Lambda_k \cdot \Lambda_u$ et $\Lambda' = \Lambda_k \cdot \Lambda'_u$ sont respectivement la matrice exacte et la matrice estimée des paramètres intrinsèques de la primitive. Un cas singulier se présente lorsque $K'_u \neq K_u$ et $\Lambda'_u \neq \Lambda_u$. À partir de l'équation (6) et de : $X = \gamma K_k \cdot K_u \cdot R \cdot \Lambda_u \cdot \Lambda_k$, on obtient :

$$R^T \cdot K_u^T \cdot K'_u{}^{-T} \cdot K'_u{}^{-1} \cdot K_u \cdot R \sim \Lambda_u^{-T} \cdot \Lambda'^T_u \cdot \Lambda'_u \cdot \Lambda_u^{-1}. \quad (9)$$

Il y a donc des singularités lorsque les matrices : $\omega'' = K_u^T \cdot K'_u{}^{-T} \cdot K'_u{}^{-1} \cdot K_u$ et $\mu'' = \Lambda_u^{-T} \cdot \Lambda'^T_u \cdot \Lambda'_u \cdot \Lambda_u^{-1}$ sont différentes de l'identité. Les matrices ω'' et μ'' ont des formes dépendant des informations connues sur les paramètres intrinsèques (le tableau 1 illustre les différents cas pour ω''), mais non de leurs valeurs. La configuration est dégénérée pour le calibrage si et seulement si l'orientation relative de la caméra et du parallélépipède est telle qu'il existe des matrices ω'' et μ'' différentes de l'identité vérifiant :

$$R^T \cdot \omega'' \cdot R \sim \mu''. \quad (10)$$

À partir de cette relation, il est possible d'obtenir les orientations relatives singulières pour le calibrage via des calculs formels. Le tableau 2 récapitule les singularités pour les cas pratiques le plus fréquents. Les singularités sont décrites de manière géométrique, en donnant l'orientation relative de la primitive par rapport à la caméra.

Paramètres intr. connus	(A) Aucun	(B) τ	(C) u_0, v_0	(D) τ, u_0, v_0
Forme de ω''	$\begin{pmatrix} a & 0 & d \\ 0 & b & e \\ d & e & c \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & d \\ 0 & 1 & e \\ d & e & c \end{pmatrix}$	$\begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & c \end{pmatrix}$

TAB. 1- La structure de la matrice ω'' (à un facteur d'échelle près) en fonction des connaissances sur les paramètres intrinsèques de la caméra. La structure de la matrice μ'' est analogue.

Trois angles droits, deux rapports de longueurs. Dans ce cas, toute la structure euclidienne du parallélépipède est connue (à un facteur d'échelle près). Ce dernier peut alors être considéré comme une mire de calibrage classique. Les singularités génériques pour le calibrage avec une mire sont décrites dans [1].

Connaissances	Configurations singulières
A-3-1	v est orthogonale à l'axe x ou y de la caméra
B-3-1	v est parallèle à l'axe optique de la caméra
C-3-1	v est parallèle à un des axes de la caméra
D-3-1	v est parallèle à l'axe optique
A-3-0	toute configuration (3 contraintes pour 4 paramètres de la caméra)
B-3-0	une des arêtes est parallèle au plan d'image
C-3-0	une des arêtes est parallèle à un des axes de la caméra
D-3-0	une des arêtes est parallèle à l'axe optique
A-2-2	impossible à décrire
B-2-2	$v \parallel$ au plan image et $w \parallel$ à l'axe optique ou plan image
C-2-2	$v \parallel$ à l'axe x ou y de la caméra et w forme un angle de 45° avec le plan image
D-2-2	$v \parallel$ à z et $w \parallel$ au plan image et forme un angle de 45° avec les axes x et y de la caméra aucune singularité
A-2-1	toute configuration (3 contraintes pour 4 paramètres de la caméra)
B-2-1	v est parallèle au plan d'image
C-2-1	v est parallèle à un des axes de la caméra ; v et w sont orthogonales à l'axe x ou y de la caméra ; v et w sont parallèles au plan d'image ;
D-2-1	v et w sont parallèles au plan d'image
A-2-0	toute configuration (2 contraintes pour 4 paramètres de la caméra)
B-2-0	toute configuration (2 contraintes pour 3 paramètres de la caméra)
C-2-0	v orthogonale à l'axe x ou y de la caméra ou \parallel plan image
D-2-0	v parallèle au plan image ou à l'axe optique

TAB. 2- Les orientations relatives provoquant des singularités pour différentes connaissances sur les paramètres intrinsèques. Chaque connaissance est de la forme X - Y - Z , où X fait référence au tableau 1 et Y, Z sont respectivement le nombre d'angles droits et le nombre de rapports de longueurs connus. Plus d'explications sont données dans la texte.

Trois angles droits, un rapport de longueurs (cas *-3-1 dans le tableau 2). Dans le tableau 2, v représente une des quatre arêtes parallèles de parallélépipède de rapport de longueurs inconnu.

Deux angles droits (cas *-2-* dans le tableau 2). Dans ce cas, le parallélépipède peut être considéré comme une primitive construite autour de deux rectangles, partageant une arête v . Une des deux arêtes du rectangle non parallèle à v est dénommée w .

3.5 Résultats expérimentaux

Afin de tester la sensibilité de la méthode proposée en présence de bruit, nous avons effectué plusieurs séries de tests sur des données synthétiques. Ces données sont constituées des projections des sommets d'un parallélépipède et des paramètres intrinsèques connus de la caméra. Nous avons aussi effectué des tests pour plusieurs orientations relatives de la caméra et du parallélépipède, tests qui viennent confirmer l'existence des singularités décrites dans le paragraphe précédent. Le parallélépipède utilisé a pour paramètres : $\{\theta_{12}, \theta_{13}, \theta_{23}\} = \{\pi/2, \pi/3, \pi/2\}$ et $\{l_1, l_2, l_3\} = \{120, 250, 130\}$; et la caméra : $\alpha_u = 500, \alpha_v = 800, u_0 = 256, v_0 = 256$ pour une taille d'image de 512×512 . En général, le calibrage a été effectué en utilisant la connaissance d'angles droits. Chaque scénario a été testé avec des amplitudes de bruit différentes (300 pour chaque scénario). Les figures de ce paragraphe montrent la moyenne et l'écart-type des résultats de calibrage pour α_v , les erreurs pour α_u étant du même ordre de grandeur.

Les expériences montrent que l'erreur pour α_v augmente linéairement avec l'amplitude du bruit des positions image des sommets du parallélépipède. Les figures 2-(a) et 2-(b) présentent les erreurs de calibrage pour un bruit uniforme d'amplitude de deux pixels et des orientations différentes du parallélépipède. Dans cette expérience, un parallélépipède possédant deux angles droits et penché de 15° par rapport au plan image subit des rotations autour de l'axe vertical (la situation est illustrée sur la figure 2-(a)). La figure 2-(b) montre que des singularités apparaissent lorsque l'axe entre les deux faces carrées du parallélépipède

est parallèle au plan image ou orthogonale à l'axe x de l'image (angles de rotation égaux à 90° et $0^\circ/180^\circ$). Cela correspond au cas C-2-0 du tableau 2.

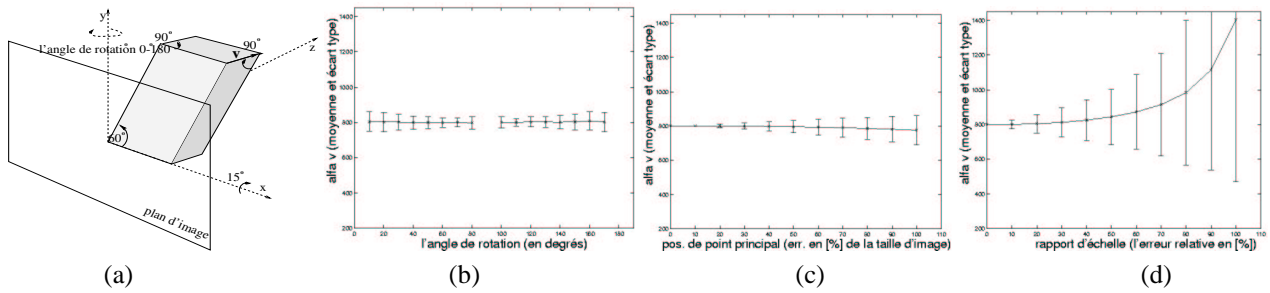


FIG. 2- (a) Le contexte de l'expérience (deux angles droits sont ici connus). (b) La moyenne et l'écart-type de α_v (la valeur exacte est 800) en fonction de l'orientation relative de la caméra et du parallélépipède. (c) La valeur de α_v en fonction d'une erreur d'hypothèse sur la position du point principal. (d) La valeur de α_v en fonction d'une erreur d'hypothèse sur le rapport d'échelle.

Les figures 2-(c) et 2-(d) montrent l'influence sur le calibrage d'hypothèses erronées sur la position du point principal et sur le rapport d'échelle. Dans la figure 2-(c), il apparaît que l'erreur sur α_v croît presque linéairement avec la distance entre le point principal exact et celui supposé. Même si on suppose que le point principal se situe sur le bord de l'image, l'erreur relative reste inférieure à 11%. Ceci confirme que l'hypothèse du point principal au centre de l'image est tout à fait raisonnable, et ce dans un contexte d'applications assez large. Par contre, la figure 2-(d) montre que l'erreur sur α_v croît de façon quadratique en fonction de l'erreur d'hypothèse sur le rapport d'échelle. En conséquence, une hypothèse sur la valeur du rapport d'échelle (fréquemment supposé égal à un) n'est réaliste que si elle repose sur une connaissance très précise de cette valeur.

4 Reconstruction 3D à partir d'une seule image

La méthode de calibrage présentée débouche naturellement sur une application simple et efficace de reconstruction de scène à partir d'une seule image. Peu d'informations et d'interactions sont en effet suffisantes pour calibrer la caméra et construire un modèle tridimensionnel – l'utilisateur doit seulement cliquer quelques points dans l'image pour définir les positions des primitives. Une telle application se révèle être un outil rapide et intuitif pour construire des modèles à partir d'images d'origines quelconques, pour lesquelles aucune information n'est disponible.

La reconstruction est indépendante du processus de calibrage décrit précédemment. L'idée est ici d'utiliser des contraintes linéaires liant les points du modèle les uns aux autres. Par exemple le fait que quatre points forment un parallélogramme, ou le fait que des points d'intérêt sont colinéaires ou coplanaires. Ces contraintes linéaires permettent de construire une structure polygonale qui inclut les parallélépipèdes utilisés pour le calibrage ainsi que des points ajoutés par la suite pour enrichir le modèle. Une description plus détaillée de la façon d'utiliser ces contraintes apparaît dans [20] (voir aussi des travaux similaires [3], [4], [5], [6], [7], [13], [16], [19]).

En résumé, l'algorithme complet de construction de modèle à partir d'une seule image se déroule de la manière suivante :

1. Les projections d'un ou plusieurs parallélépipèdes sont définies et les paramètres intrinsèques de la caméra sont déterminés.
2. Les points additionnels du modèle et les contraintes linéaires qui leurs sont associées sont définis. Toutes les contraintes forment un système linéaire. La résolution de ce système fournit les valeurs des profondeurs des points sur les rayons de projection.
3. Le modèle est créé à partir des points reconstruits et de l'image de départ (pour les textures).

4.1 Quelques résultats expérimentaux

Une interface basée sur la librairie OpenGL a été développée à partir des idées décrites dans cet article. Cette interface permet, en quelques minutes, de construire des modèles réalistes à partir de photos non calibrées et de les visualiser. Les modèles construits peuvent aussi être exportés dans le format standard VRML. Dans ce paragraphe, nous présentons deux modèles générés avec notre méthode.

4.1.1 Reconstruction d'une scène intérieure

La figure 3 présente l'image originale et les rendus du modèle construit d'une chambre. La plupart des angles dans la scène ne sont pas parfaitement droits (la déviation moyenne mesurée est de $\sim 3^\circ$). Le placard situé dans la partie centrale de l'image a servi au calibrage. Pour la reconstruction, nous avons utilisé plusieurs contraintes différentes – trois ensembles connectés de parallélogrammes (la partie centrale de l'image, la fenêtre et la poutre) ont été connectés en utilisant des contraintes de coplanarité. Les murs ont été reconstruits en utilisant des contraintes de coplanarité et de colinéarité. Le modèle final a été vérifié en mesurant des angles supposés droits. L'erreur relative maximale est ici de 1.8%.

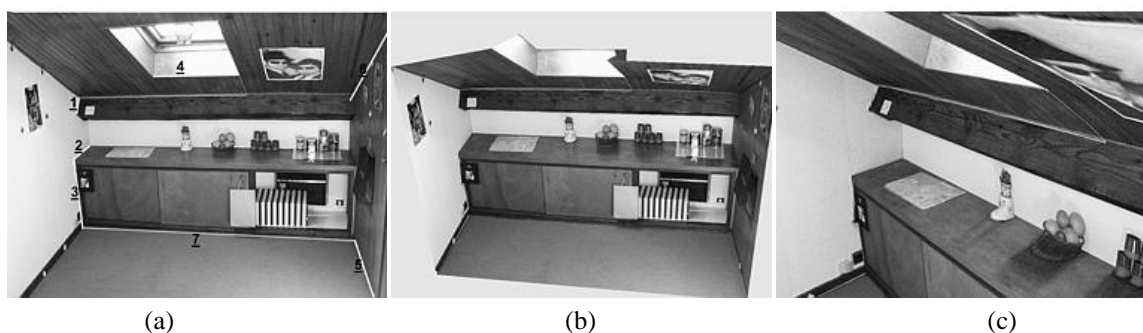


FIG. 3- (a) L'image originale. (b), (c) Images du modèle de la scène intérieure.

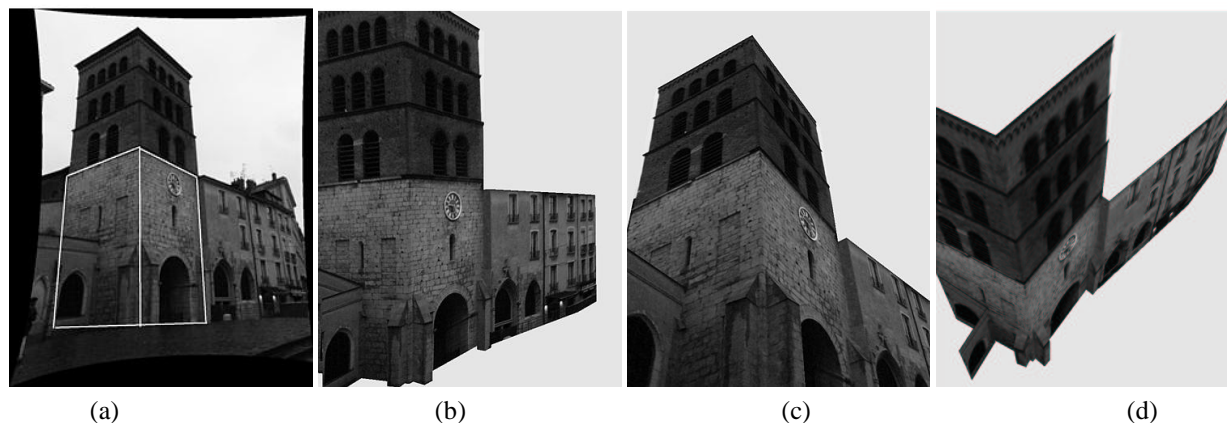


FIG. 4- (a) L'image originale. (b), (c), (d) Images du modèle de la scène extérieure.

4.1.2 Reconstruction d'une scène extérieure

La distorsion de l'image originale a tout d'abord été corrigée en utilisant des paramètres de distorsion de référence. La figure 4 montre l'image après correction et la primitive utilisée pour le calibrage. Le rapport d'échelle a été calibré avec 1.8% d'erreur relative et α_v avec 1.9%, ceci par rapport à un calibrage établi à l'aide d'une mire. Les angles de la tour sont reconstruits avec une erreur relative maximale de 5%.

5 Conclusion et perspectives

Nous avons présenté une méthode de calibrage de caméra et de reconstruction de scène basée sur les parallélépipèdes. Nous avons montré que : (i) les caractéristiques intrinsèques d'un parallélépipède et d'une caméra perspective sont duaux ; (ii) les connaissances d'angles droits et de rapports de longueurs dans la scène ainsi que des informations sur les paramètres intrinsèques de la caméra peuvent être utilisées simultanément dans un système linéaire ; (iii) une hypothèse sur la position du point principal est plus robuste que celle sur le rapport d'échelle ; (iv) un calibrage correct peut être obtenu en utilisant un ou plusieurs parallélépipèdes avec seulement deux angles droits (ce qui est souvent le cas dans les environnements humains). Une étude sur les singularités de la méthode a été effectuée, délimitant ainsi le contexte d'application de manière précise. L'approche proposée dans cet article a été illustrée par des exemples probants de reconstruction à partir d'une seule image.

Plusieurs extensions de cette approche sont possibles. Dans un premier temps, il serait utile d'automatiser la procédure de recherche des primitives dans l'image. Nous envisageons par la suite d'étendre la méthode à plusieurs images.

Références

- [1] T. Buchanan, "The Twisted Cubic and Camera Calibration," *CVGIP*, 42(1), pp. 130–132, 1988.
- [2] B. Caprile et V. Torre, "Using Vanishing Points for Camera Calibration," *IJCV*, 4(2), pp. 127–140, 1990.
- [3] Sparr, G., "Depth Computations From Polyhedral Images," *ECCV*, pp. 378–386, 1992.
- [4] C.S. Chen, C.k. Yu, et Y.P. Hung, "New Calibration-free Approach for Augmented Reality Based on Parameterized Cuboid Structure," *ICCV*, pp. 30–37, 1999.
- [5] R. Cipolla et E. Boyer, "3D model acquisition from uncalibrated images," *IAPR Workshop on Machine Vision Applications*, pp. 559–568, 1998.
- [6] A. Criminisi, I. Reid, et A. Zisserman, "Single View Metrology," *ICCV*, pp. 434–441, 1999.
- [7] P.E. Debevec, C.J. Taylor, et J. Malik, "Modeling and Rendering Architecture from Photographs: A Hybrid Geometry- and Image-Based Approach," *SIGGRAPH*, 1996.
- [8] R.I. Hartley et A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, 2000.
- [9] Y. Horry, K. Anjyo, et K. Arai, "Tour Into the Picture," *SIGGRAPH*, 1997.
- [10] K. Kanatani, *Statistical Optimization for Geometric Computation: Theory and Practice*, Lecture Note, 1995.
- [11] D. Liebowitz et A. Zisserman, "Metric Rectification for Perspective Images of Planes," *CVPR*, pp. 482–488, 1998.
- [12] E. Malis et R. Cipolla, "Multi-View Constraints Between Collineations: Application to Self-Calibration from Unknown Planar Structures," *ECCV*, pp. 610–624, 2000.
- [13] P.F. Sturm et S.J. Maybank, "A Method for Interactive 3D Reconstruction of Piecewise Planar Objects from Single Images," *BMVC*, pp. 265–274, 1999.
- [14] W.H. Press et B.P. Flannery and S.A. Teukolsky and W.T. Vetterling, *Numerical Recipes in C, The Art of Scientific Computing, Second Edition*, Cambridge University Press, 1992.
- [15] P.F. Sturm et S.J. Maybank, "On Plane-Based Camera Calibration: A General Algorithm, Singularities, Applications," *CVPR*, pp. 432–437, 1999.
- [16] P. Sturm, "A Method for 3D Reconstruction of Piecewise Planar Objects from Single Panoramic Images," *CVPR*, pp. 119–126, 2000.
- [17] B. Triggs, "Autocalibration from Planar Scenes," *ECCV*, pp. 89–105, 1998.
- [18] Z. Zhang, "A flexible new technique for camera calibration," *IEEE-PAMI*, 22(11), pp. 1330–1334, 2000.
- [19] D. Liebowitz et A. Zisserman, "Combining Scene and Auto-calibration Constraints", *ICCV*, pp.293-300, 1999.
- [20] M. Wilczkowiak, E. Boyer, P. Sturm, "Camera Calibration and 3D Reconstruction from Single Images Using Parallelepipeds", *ICCV*, 2001.