

Triangulation projective contrainte par multi-coplanarité

Adrien Bartoli et Peter Sturm

INRIA Rhône-Alpes, 655, avenue de l'Europe, 38334 Saint Ismier Cedex.

Adrien.Bartoli@inria.fr—Peter.Sturm@inria.fr

Résumé

Cet article traite du problème de la reconstruction et de l'estimation du mouvement entre les caméras à partir de deux vues d'une scène rigide. Plus particulièrement, nous traitons le cas où la scène contient des plans, c'est-à-dire des ensembles de points coplanaires ou multi-coplanaires (points sur plusieurs plans), qui constituent des contraintes géométriques très fortes. La plupart des travaux existants ne les exploitent que d'une manière sous-optimale. Une approche typique est d'estimer une reconstruction isolée de points, d'ajuster des plans et éventuellement de corriger la position 3D des points afin de les rendre coplanaires. Dans cet article, nous présentons une méthode permettant d'estimer de façon conjointe et optimale (au sens du maximum de vraisemblance) la structure de la scène et le mouvement entre les caméras : le résultat est une structure minimisant l'erreur de reprojexion tout en satisfaisant exactement les contraintes géométriques. Pour ce faire, la structure est paramétrée de façon minimale par des entités 2D ou 3D. Des résultats expérimentaux montrent que les résultats de reconstruction obtenus sont de qualité supérieure à ceux d'autres méthodes, notamment celles basées sur une reconstruction individuelle de points, ceci même dans les cas où les plans de la scène sont imparfaits.

Mots clés : reconstruction projective, multi-coplanarité, stéréovision.

1 Introduction

L'estimation de la structure ou reconstruction à partir d'images constitue un des aspects les plus importants du domaine de la photogrammétrie et de la vision par ordinateur. Le cas des scènes planes par morceaux est particulièrement important de par leur nombre élevé en environnement humain.

Dans cet article, nous proposons une approche projective optimale pour l'estimation de la structure à partir de deux vues d'une scène plane par morceaux, qui a la flexibilité de travailler à partir d'images sans ou avec des données de calibrage partielles.

La plupart des méthodes existantes [1, 4, 11, 12, 14] résolvent ce problème de manière sous-optimale via une reconstruction isolée de primitives (points ou droites) dans un cadre calibré. Les contraintes géométriques de la structure, en d'autres termes sa planarité par morceaux, traduite par l'appartenance de points à des plans, ne sont pas exactement satisfaites mais approchées.

En effet, seule une représentation algébrique prenant explicitement en compte ces contraintes permet d'obtenir une structure les satisfaisant [2]. Ceci se traduit par le fait que seuls les ddl (degrés

de liberté) essentiels de la structure sont reflétés par une telle paramétrisation (par exemple, un point appartenant à un plan voit son nombre de ddl réduit de 3 à 2).

D'autre part, un cadre projectif impose d'utiliser des critères au niveau image (il n'existe pas de métrique dans un tel espace), ce qui ajoute une difficulté car la structure est représentée indirectement (par exemple, un plan à 3 ddl est souvent représenté par une matrice d'homographie à 9 ddl) entraînant une sur-paramétrisation.

Travaux antérieurs Un estimateur au maximum de vraisemblance consiste, dans notre cas, à minimiser une fonction de coût donnée par le résidu de la reprojexion (distance euclidienne entre les points donnés et les points reprojétés) par rapport à tous les paramètres inconnus du modèle. Cette technique, aussi appelée la méthode directe [13], est optimale sous l'hypothèse d'un bruit gaussien centré et indépendant sur les points image donnés. Les paramètres du modèle pouvant être inclus dans cette optimisation sont, en particulier, la structure $3D$, les matrices de projection et le cas écheant, les données de calibrage. Dans cet article, le terme optimal sera utilisé dans le sens du maximum de vraisemblance et sous les hypothèses décrites ci-dessus.

La fonction de coût est, dans notre cas, non linéaire. La phase d'optimisation est donc itérative et nous la conduisons via l'algorithme de Levenberg-Marquardt [9].

La méthode proposée dans [12] permet d'obtenir la structure $3D$ sous forme d'un ensemble de facettes planes. Le critère proposé ne donne pas une structure optimale car les contraintes de jointure entre plans ne sont pas prises en compte, et il ne correspond pas à celui de la méthode directe. De manière similaire, les résultats obtenus dans [1, 4] sont sous-optimaux.

La solution établie dans [14] consiste à modéliser la structure comme un ensemble de points $3D$ et à minimiser le résidu de la reprojexion basé sur les homographies planes induites par les plans de la scène. La coplanarité par morceaux des points n'y est pas forcément vérifiée.

La prise en compte de cette contrainte est proposée dans [11] où les auteurs ajoutent à leur critère (de même type que celui de [14]) des termes de pénalité fortement pondérés censés faire tendre la structure vers une configuration plane par morceaux. Cette approche présente deux inconvénients : la planarité par morceaux n'est qu'approchée et les équations additionnelles peuvent entraîner des instabilités numériques lors de la phase d'optimisation.

Dans ces travaux, en aucun cas la planarité par morceaux n'est prise en compte explicitement par une paramétrisation adéquate de la structure (cette dernière y est paramétrée comme générale, sans contrainte géométrique). De plus, les images utilisées sont toutes supposées être accompagnées de données de calibrage (complètes [1, 4, 11, 12] ou sans les focales [14]).

Dans [2], une paramétrisation de la structure permet de prendre en compte explicitement l'appartenance d'un point à un plan dans un cadre projectif et de réaliser une estimation optimale de cette dernière. Cependant, l'appartenance éventuelle d'un point à plusieurs plans n'est pas gérée.

Contributions, organisation Le résultat principal donné dans cet article est une paramétrisation consistante exprimant la structure $3D$ de la scène et les contraintes géométriques qui lui sont associées. Cette paramétrisation est entièrement décrite au niveau image, à l'exception des points n'appartenant à aucun plan qui sont représentés dans l'espace. Les contraintes géométriques abordées sont la multicoplanarité, c'est-à-dire l'appartenance d'un point à plusieurs plans. Nous étudions les cas où un point appartient à un nombre de plans allant de zéro à trois. Les cas où un point est contenu par plus de plans, étant très rares, présentent peu d'intérêt.

La paramétrisation de la structure est présentée dans la section 2. L'estimateur au maximum de vraisemblance est ensuite déduit, ainsi que l'initialisation de la structure, dans la section 3. Finale-

ment, nous donnons nos résultats expérimentaux et nos conclusions et perspectives dans les sections 4 et 5 respectivement.

Préliminaires Le mouvement entre deux caméras peut être décrit sous sa forme projective par la géométrie épipolaire contenue dans la matrice fondamentale F [7]. Cette dernière permet de définir une base projective canonique pour la reconstruction, donnée en terme des matrices de projection canoniques définies pour chacune des deux images par $P \sim (I|0)$ et $P' \sim ([e']_{\wedge} F | e')$ respectivement. L'épipôle droit e' est défini par $F^T e' = 0$, $[\cdot]_{\wedge}$ représente la matrice anti-symétrique associée au produit vectoriel (telle que $[x]_{\wedge} y = x \wedge y$) et \sim l'égalité à un facteur multiplicatif non nul près. Les entités géométriques (points, droites, plans, etc.) sont désignées par des caractères italiques et leurs vecteurs de coordonnées dans une certaine base par les mêmes caractères droits et en gras, par exemple x et \mathbf{x} pour un point.

Un plan π dont l'équation est $\pi^T \sim (\bar{\pi}^T, 1)$ dans la base canonique induit une homographie plane H_{π} entre les deux images donnée par [8] :

$$H_{\pi} \sim [e']_{\wedge} F - e' \bar{\pi}^T, \quad (1)$$

qui s'applique aux projections x et x' d'un point X appartenant à π via :

$$x' \sim H_{\pi} x. \quad (2)$$

L'équation réduite $\bar{\pi}$ est caractéristique d'un plan π .

La projection l dans la première image de la droite formée par l'intersection de deux plans π et π' est donnée par :

$$l \sim \bar{\pi} - \bar{\pi}'. \quad (3)$$

Cette égalité peut être démontrée en utilisant l'expression d'un point X appartenant à un plan π [2] : $X^T \sim (x^T, -x^T \bar{\pi})$. Si $X \in \pi$ et $X \in \pi'$, nous obtenons $x^T \bar{\pi} = x^T \bar{\pi}'$ puis $x^T (\bar{\pi} - \bar{\pi}') = 0$.

2 Paramétrisation de la structure

Dans cette section, nous proposons une représentation algébrique de la structure, que l'on notera ν , satisfaisant les contraintes géométriques de la scène (appartenance des points à un ou plusieurs plans).

Nous donnons tout d'abord une représentation des plans puis des points, selon leur appartenance à aucun, un seul, deux ou trois des plans modélisés.

2.1 Plans

Un plan π est modélisé par son équation réduite $\bar{\pi}$ dans la base canonique. Etant un vecteur-3 non homogène, $\bar{\pi}$ possède bien les 3 ddl requis. De plus, l'homographie plane H_{π} se déduit directement via l'équation (1).

2.2 Points

Un point X a un nombre différent de ddl selon le nombre de plans modélisés auxquels il appartient.

Points n'appartenant à aucun plan Le point X n'est pas contraint et garde donc ses 3 ddl. Nous utilisons une représentation directe dans la base projective canonique par un vecteur-4 homogène \mathbf{X} . Comme toute entité homogène, ce vecteur est défini à un facteur multiplicatif non nul près. Cette ambiguïté est levée en utilisant la contrainte de norme $\|\mathbf{X}\|^2 = 1$. Cette représentation possède bien les 3 ddl requis. Les points image x et x' correspondants sont obtenus par projection de X . Une paramétrisation équivalente au niveau image est possible en utilisant la contrainte $\mathbf{x}^T \mathbf{F} \mathbf{x} = 0$, mais elle est algébriquement plus difficile à manipuler que la simple contrainte de norme donnée ci-dessus.

Points n'appartenant qu'à un seul plan Le point X contraint d'appartenir au plan π a 2 ddl. Ses deux projections sont reliées par l'homographie plane H_π selon l'équation (2). Par conséquent, représenter X par un point dans une seule image, x par exemple, permet d'obtenir le correspondant dans l'autre image et possède bien les 2 ddl requis pour X .

Points appartenant à deux plans Le point X appartenant à la droite d'intersection des plans π et π' n'a qu'un seul ddl. Comme dans le cas précédent, cherchons une représentation de X au niveau

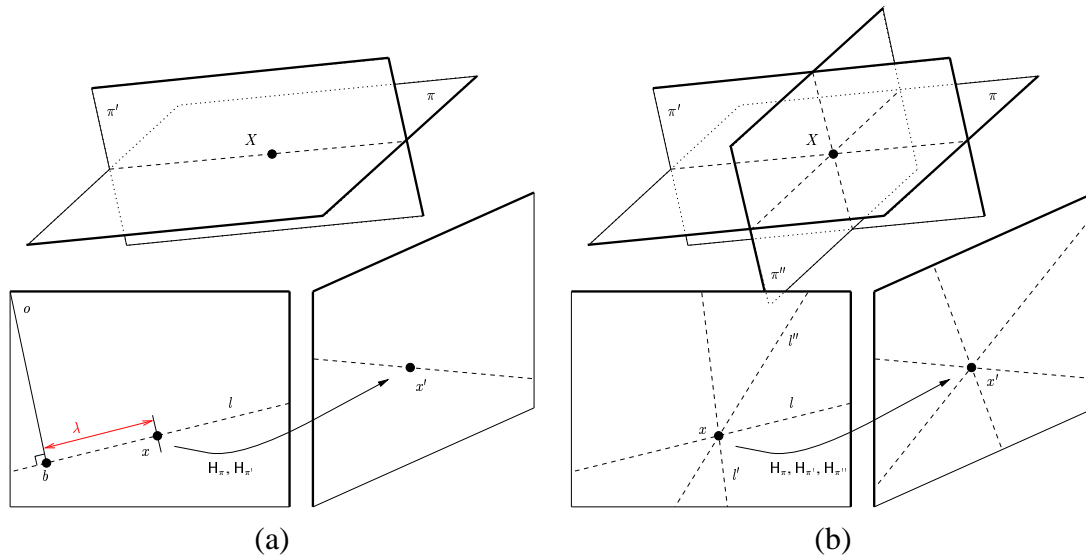


FIG. 1 : paramétrisation d'un point appartenant à deux (a) ou trois plans (b).

image. Le point image x permet via une des deux homographies planes H_π ou $H_{\pi'}$ d'obtenir x' selon l'équation (2).

Le problème se réduit donc à exprimer le point image x par un seul paramètre, sachant qu'il appartient à une droite l , la projection de l'intersection des plans π et π' , donnée par l'équation (3).

Si l'on fixe arbitrairement un point b de l , nous pouvons alors représenter tout point image x de l par sa mesure algébrique λ avec b . La figure 1a illustre cette construction. Établissons la correspondance entre λ et x .

Si l'on définit la direction normalisée de l par $\Delta_l = \frac{l \wedge l_\infty}{\|l \wedge l_\infty\|}$ (la direction est un point à l'infini), le point x est obtenu en ajoutant Δ_l pondérée par λ à b :

$$\mathbf{x} \sim \begin{pmatrix} \bar{\mathbf{b}} \\ 1 \end{pmatrix} + \lambda \Delta_l \sim \begin{pmatrix} \bar{\mathbf{b}} + \frac{\lambda}{l_1^2 + l_2^2} \begin{pmatrix} l_2 \\ -l_1 \end{pmatrix} \\ 1 \end{pmatrix}. \quad (4)$$

$X \in$	param.	commentaires	contraintes	\mathbf{x}	\mathbf{x}'
\mathbb{P}^3	\mathbf{X}	vecteur-4 homogène	$\ \mathbf{X}\ ^2 = 1$	\mathbf{PX}	$\mathbf{P}'\mathbf{X}$
π	\mathbf{x}	vecteur-2 non homogène	\emptyset	\mathbf{x}	$\mathbf{H}_\pi \mathbf{x}$
π, π'	λ	mesure algébrique sur l	\emptyset	éqn. (4)	$\mathbf{H}_\pi \mathbf{x}$
π, π', π''	\emptyset	\cap des 3 plans	\emptyset	éqn. (5)	$\mathbf{H}_\pi \mathbf{x}$

TAB. 1 : paramétrisation des points de la structure selon le nombre de plans (entre 0 et 3) auxquels ils appartiennent.

Le point b peut être choisi comme la projection orthogonale sur l du centre o du système de coordonnées image. L'expression de b est alors donnée par :

$$\mathbf{b} \sim \begin{pmatrix} \bar{\mathbf{o}} - \frac{\mathbf{o}^T \bar{\mathbf{l}}}{\|\bar{\mathbf{l}}\|^2} \bar{\mathbf{l}} \\ 1 \end{pmatrix} \sim \begin{pmatrix} -\frac{l_3}{l_1^2 + l_2^2} \begin{pmatrix} l_1 \\ l_2 \end{pmatrix} \\ 1 \end{pmatrix} \text{ avec } \mathbf{o}^T \sim (\bar{\mathbf{o}}, 1)^T \sim (0, 0, 1)^T.$$

Points appartenant à trois plans Le point X appartient à trois des plans modélisés, π , π' et π'' . L'intersection de trois plans étant en général un point, X n'a aucun ddl (figure 1b). Le problème revient donc à trouver une expression donnant la projection de X dans une des deux images, sa position dans l'autre image étant alors donnée via une des homographies planes \mathbf{H}_π , $\mathbf{H}_{\pi'}$ ou $\mathbf{H}_{\pi''}$. Comme pour les cas précédents, cherchons une expression pour le point image x .

Soient l , l' et l'' les droites projection de $\pi \cap \pi'$, $\pi' \cap \pi''$ et $\pi'' \cap \pi$. Il est facile de montrer que ces trois droites s'intersectent en un point unique, x . En effet, \mathbf{x} est le noyau de $\mathbf{A} \sim (1|l'|l'')^T$ et le déterminant de \mathbf{A} est nul car, d'après l'équation (3), toute ligne de \mathbf{A} est combinaison linéaire des deux autres, donc x est toujours défini. Son expression est obtenue en utilisant l'intersection de deux des trois droites l , l' et l'' :

$$\mathbf{x} \sim (\bar{\pi} \wedge \bar{\pi}') + (\bar{\pi}' \wedge \bar{\pi}'') + (\bar{\pi}'' \wedge \bar{\pi}). \quad (5)$$

Elle est identique quelle que soit la paire de droites choisie parmi l , l' et l'' , ce que reflète bien l'existence du noyau de \mathbf{A} .

2.3 Résumé de la paramétrisation

Nous résumons ici les résultats obtenus pour la paramétrisation de la structure d'une scène contenant des plans et des points sur $[0 \dots 3]$ plans. Un plan π est représenté par un vecteur-3 non homogène $\bar{\pi}$ tel que $\pi^T \sim (\bar{\pi}^T, 1)$. L'homographie plane induite par π entre les deux images est donnée par $\mathbf{H}_\pi \sim [\mathbf{e}']_\wedge \mathbf{F} - \mathbf{e}' \bar{\pi}^T$. La représentation des points est donnée dans le tableau 1.

La paramétrisation de la structure est réunie dans un ensemble $\nu = \{\bar{\pi}\} \cup \{\mathbf{X}\} \cup \{\mathbf{x}\} \cup \{\lambda\} \cup \emptyset$ contenant respectivement la paramétrisation des plans et des points n'appartenant à aucun plan, appartenant à un, deux ou trois plans.

3 Structure optimale

Dans cette section, nous utilisons la paramétrisation définie précédemment afin d'obtenir l'estimateur donnant la structure optimale. Nous proposons tout d'abord une méthode d'initialisation directe de la structure, puis nous donnons le critère correspondant au maximum de vraisemblance.

3.1 Initialisation

La paramétrisation des plans est initialisée via l'estimation des homographies planes correspondantes. Ceci est possible car la géométrie épipolaire peut être estimée par un algorithme direct [6] à partir de correspondances de points. Les points nécessitent une initialisation spécifique dépendant du nombre de plans auxquels ils appartiennent.

Plans Chaque correspondance $x \leftrightarrow x'$ liée à un point 3D X appartenant à un plan π est soumise à l'équation (2) où l'homographie plane H_π induite par π peut être écrite en fonction de $\bar{\pi}$, l'équation réduite de π selon l'équation (1), ce qui donne $\mathbf{x}' \sim ([\mathbf{e}']_\wedge F - \mathbf{e}'\bar{\pi}^\top)\mathbf{x}$. L'égalité \sim à un facteur multiplicatif près de deux vecteurs-3 équivaut à annuler le produit vectoriel de ces deux vecteurs [6]. En utilisant cette propriété, et après quelques manipulations algébriques mineures, nous obtenons $[\mathbf{x}']_\wedge \mathbf{e}'\mathbf{x}'^\top \bar{\pi} = [\mathbf{x}']_\wedge [\mathbf{e}']_\wedge F\mathbf{x}$.

Chaque correspondance de points donne une égalité de cette forme. Ces équations peuvent alors être réarrangées en un système linéaire en $\bar{\pi}$. Cette méthode est applicable à partir de trois correspondances de points, ce qui correspond bien au fait qu'un plan est défini par trois points. La fonction de coût correspondant à cet estimateur est donnée par $\sum_i \|[\mathbf{x}'_\wedge H\mathbf{x}_i]\|^2$ où $H \sim [\mathbf{e}']_\wedge F - \mathbf{e}'\bar{\pi}^\top$.

Un estimateur d'homographie plane basé sur une approche quasi-linéaire est proposé dans [2]. Le gain de précision qu'il pourrait apporter n'est pas nécessaire, ceci à cause de l'optimisation non linéaire qui suit.

Point n'appartenant à aucun plan L'initialisation d'un tel point correspond à sa triangulation dans la base projective canonique. Une solution optimale et directe est donnée dans [5].

Point n'appartenant qu'à un seul plan La solution optimale pour de tels points nécessite de résoudre une équation algébrique de degré 8 [10], ce qui est coûteux d'une part et instable d'autre part. C'est pourquoi nous optons pour une initialisation consistant à prendre le point x tel qu'il est. Le point $H_\pi x$ sera en général différent du point mesuré x' . Ceci introduit un biais, mais qui en général ne sera pas très important et qui sera levé par l'optimisation non linéaire ultérieure.

Point appartenant à deux plans Une solution optimale nécessiterait une méthode du même type que [5], et plus précisément, une équation algébrique de degré 4 devrait être résolue. Nous optons pour une initialisation consistant à projeter le point x sur l , la projection de la droite d'intersection des deux plans contenant X .

Point appartenant à trois plans Nous avons vu qu'un tel point n'a aucun ddl. Il ne nécessite donc pas d'initialisation.

3.2 Fonction de coût

L'estimation consiste à optimiser une fonction de coût notée \mathcal{C} correspondant à l'approche directe [13]. Les paramètres inclus dans cette optimisation sont ceux de la structure de la scène. La valeur de la fonction de coût dépend des points image mesurés $x_i \leftrightarrow x'_i$ ainsi que des points image $\hat{x}_i \leftrightarrow \hat{x}'_i$ de la structure estimée. Elle est définie par $\mathcal{C}(\nu) = \sum_i (d^2(x_i, \hat{x}_i) + d^2(x'_i, \hat{x}'_i))$, où $d(\cdot, \cdot)$ est la distance euclidienne.

Les points $\hat{x}_i \leftrightarrow \hat{x}'_i$ sont extraits de la paramétrisation $\hat{\nu}$ de la structure (réunissant les équations des plans et tous les paramètres des points, selon le tableau 1), permettant ainsi de satisfaire les contraintes géométriques. La structure optimale $\hat{\nu}$ est alors donnée par la minimisation de $\mathcal{C} : \hat{\nu} = \operatorname{argmin}_{\nu} \mathcal{C}(\nu)$ via l'algorithme d'optimisation non linéaire de Levenberg-Marquardt [9].

Un exemple de fonction de coût est donné, pour un plan π et un point X sur ce plan par $\mathcal{C}(\{\hat{\pi}, \hat{\mathbf{x}}\}) = d^2(\mathbf{x}, \hat{\mathbf{x}}) + d^2(\mathbf{x}', ([\mathbf{e}']_{\wedge} \mathbf{F} + \mathbf{e}' \hat{\pi}^T) \hat{\mathbf{x}})$.

Estimation conjointe du mouvement entre les caméras Le mouvement entre les caméras, décrit par la géométrie épipolaire peut n'avoir été donné que sous forme d'une estimation sous-optimale (même son estimation au maximum de vraisemblance est sous-optimale tant que la structure plane par morceaux de la scène n'est pas explicitement prise en compte). Dans ce cas, on peut obtenir sa version optimale, ainsi que celle de la structure, en incluant ses paramètres à la phase d'optimisation en utilisant une paramétrisation de la matrice fondamentale comme celle décrite dans [2]. La fonction de coût utilisée est identique à \mathcal{C} . Cet estimateur est expérimenté dans la section suivante.

4 Résultats expérimentaux

4.1 Données synthétiques

Dans cette section, nous comparons les résultats obtenus avec notre paramétrisation à d'autres méthodes existantes. Le banc de test consiste en un cube d'un mètre de côté à différentes distances d'une paire de caméras. Des points sont générés sur le cube selon la répartition suivante : 50 points sur chaque face, 10 points sur chaque arête et 1 point par sommet. Ces points sont projetés sur les deux caméras et un bruit gaussien centré est ajouté aux points image. La qualité d'une reconstruction X_i est mesurée par le résidu 3D de la distance euclidienne avec la vraie structure \underline{X}_i : $E_3 = \sqrt{\frac{1}{n} \sum_i d^2(\mathbf{H}_3 \mathbf{X}_i, \underline{\mathbf{X}}_i)}$, où \mathbf{H}_3 est une homographie 3D estimée par minimisation non linéaire de E_3 et n le nombre total de points.

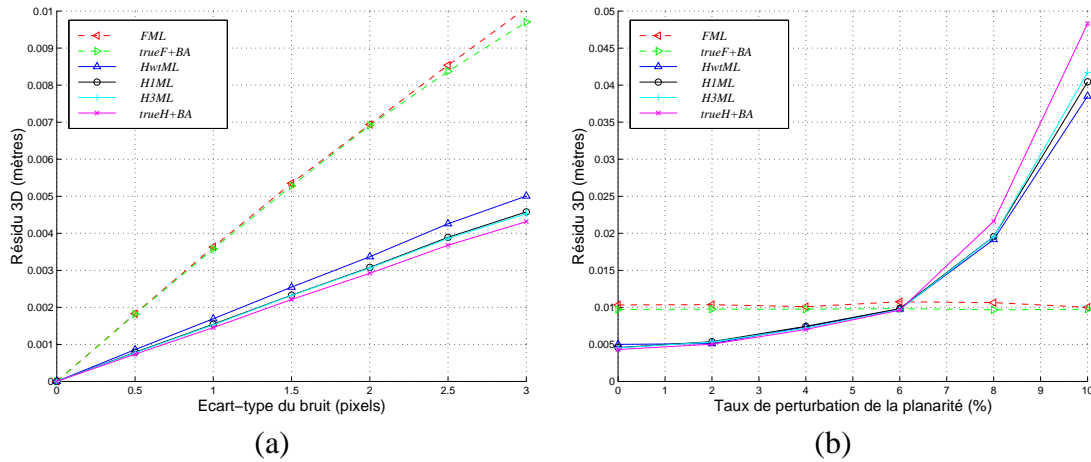


FIG. 2 : comparaison des différentes méthodes par le résidu 3D E_3 , pour des points parfaitement (a) ou approximativement coplanaires (b).

La première expérimentation (figure 2a), où les points 3D simulés sont parfaitement coplanaires, permet de conclure que les méthodes *FML* (reconstruction au maximum de vraisemblance) et *trueF*

(reconstruction optimale basée sur la géométrie épipolaire exacte) reposant uniquement sur une reconstruction individuelle de points donnent un résultat de qualité inférieure aux méthodes prenant la structure particulière de la scène en compte. Par ordre décroissant de qualité, *HwtML* décrite dans [11] prend en compte l'appartenance d'un point à un plan par des équations additionnelles. *H1ML*, proposée dans [2] prend en compte explicitement la coplanarité et *H3ML*, décrite dans cet article, la multi-coplanarité. Finalement, *trueH+BA* utilise les vrais plans et la multi-coplanarité.

La seconde expérimentation (figure 2b), montre le cas où les points 3D simulés sont écartés du cube par un bruit gaussien centré dans une direction normale à chaque face à laquelle ils appartiennent. Nous observons dans ce cas, qu'en dessous d'un seuil de rupture (rapport de l'écart-type de la perturbation et de la taille du cube lorsque les méthodes basées sur les plans deviennent moins performantes que celles basées sur des points seulement), 6% dans ce cas, les méthodes basées sur une structure plane par morceaux produisent de meilleurs résultats que celles basées sur des reconstructions individuelles de points, même si la scène n'est pas parfaitement plane. De plus, on observe que plus une méthode est contrainte, plus elle se dégrade au-delà du seuil de rupture (par exemple, *H3ML* devient moins performante que *H1ML* et *HwtML* qui sont moins contraintes), ce qui est expliqué par le fait que le modèle a priori exprimé dans les contraintes est de moins en moins valide à mesure que la perturbation de la planarité augmente. Le tableau suivant donne différentes valeurs du seuil de rupture pour différentes valeurs de la distance scène/caméras et pour différentes valeurs de l'écart-type du bruit ajouté sur les points image :

	3 m	10 m	20 m
1 pixel	0.5%	2%	4%
3 pixels	2%	6%	9%

Ces valeurs représentent des variations supérieures à celles d'une grande majorité de surfaces réelles approximativement planes.

4.2 Images réelles

Nous décrivons le processus complet de reconstruction, des images au modèle métrique 3D texturé. Cette description est illustrée en utilisant la paire d'images de la figure 3a. Nous comparons ensuite notre approche avec une reconstruction individuelle de points.

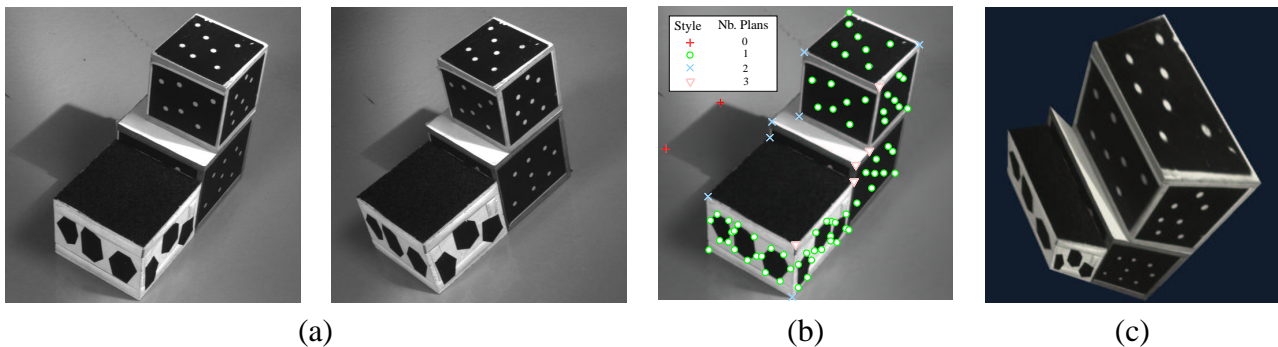


FIG. 3 : la paire d'images *cubicScene* (a), la multi-coplanarité (b) et un rendu texturé de la reconstruction projective (c).

Correspondances de points, multi-coplanarité Ces données sont établies manuellement. La figure 3b illustre la notion de multi-coplanarité. On peut aussi utiliser la méthode automatique décrite dans [1].

Textures Une délimitation polygônale de chaque facette plane est donnée sur une des deux images.

Triangulation projective L'intégralité des correspondances de points est utilisée pour estimer la géométrie épipolaire selon l'algorithme des 8 points [6]. La structure est ensuite initialisée et la fonction de coût \mathcal{C} optimisée (voir la section 3) en incluant les paramètres de la géométrie épipolaire. La figure 3c montre une vue du modèle projectif obtenu.

Structure métrique La structure métrique est obtenue par autocalibrage basé sur la matrice fondamentale [3]. Ce résultat pourrait être amélioré en utilisant la technique proposée dans [14] et basée sur la structure plane par morceaux de la scène. Des vues du modèle métrique obtenu sont sur la figure 4.

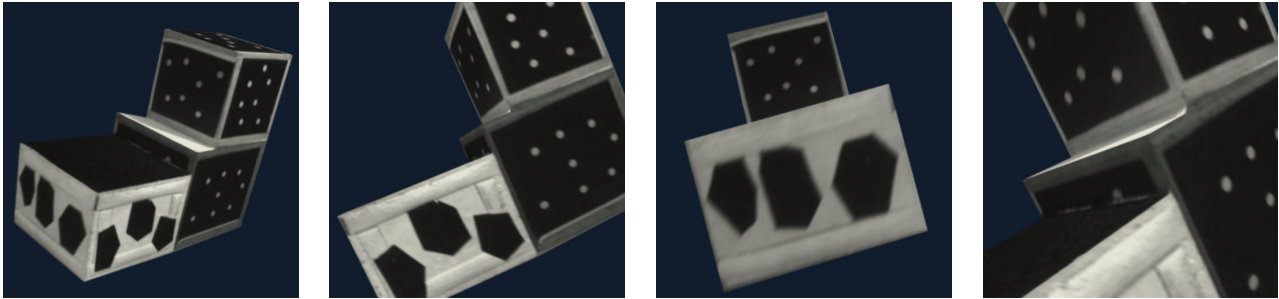
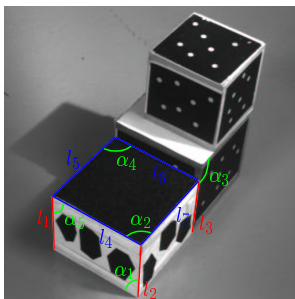


FIG. 4 : rendus texturés de la reconstruction euclidienne.

Gain de qualité Nous avons mesuré des grandeurs significatives (rapports de longueurs et angles) sur la reconstruction métrique obtenue par notre approche et sur celle obtenue par reconstruction individuelle des mêmes points image. Les valeurs obtenues (figure 5), montrent clairement le gain de précision apporté par la prise en compte de la structure plane par morceaux de la scène.



	l_1/l_2	l_1/l_3	l_4/l_5	l_4/l_6	l_4/l_7
points	1.0412	0.9534	1.1061	1.0421	1.0249
plans	1.0280	0.9655	1.0400	1.0297	1.0232

	$2\alpha_1/\pi$	$2\alpha_2/\pi$	$2\alpha_3/\pi$	$2\alpha_4/\pi$	$2\alpha_5/\pi$
points	1.1445	0.9581	1.1023	0.9530	0.9752
plans	0.9743	0.9816	0.9964	0.9857	0.9948

FIG. 5 : mesures métriques sur les reconstructions obtenues par deux méthodes différentes.

5 Conclusions et perspectives

Nous avons présenté un estimateur au maximum de vraisemblance pour la structure et le mouvement entre les caméras d'une scène plane par morceaux à partir de deux vues. La structure est représentée en tenant compte des contraintes géométriques de multi-coplanarité par une paramétrisation consistante. Les plans sont représentés par leurs équations réduites dans la base canonique et les points spécifiquement selon le nombre de plans, entre zéro et trois, auxquels ils appartiennent.

L'initialisation des équations réduites de plans est effectuée via l'estimation d'homographies planes. Des résultats expérimentaux, sur données synthétiques et images réelles montrent que la qualité de la reconstruction obtenue avec notre approche est clairement supérieure à celle donnée par des méthodes de reconstruction individuelle de points, et ceci même lorsque la scène considérée n'est pas parfaitement plane par morceaux.

Nous envisageons de concevoir un détecteur automatique de structures planes au niveau image. De plus, nous travaillons actuellement à l'extension de l'approche à un nombre quelconque de vues. Le résultat devrait être un système complet, capable de détecter et reconstruire des scènes planes par morceaux automatiquement.

Références

- [1] C. Baillard et A. Zisserman. Automatic reconstruction of piecewise planar models from multiple views. In *Proceedings of the Conference on Computer Vision and Pattern Recognition, Fort Collins, Colorado, USA*, pages 559–565. IEEE Computer Society Press, juin 1999.
- [2] A. Bartoli, P. Sturm, et R. Horaud. A projective framework for structure and motion recovery from two views of a piecewise planar scene. Rapport de recherche 4070, INRIA, Grenoble, octobre 2000.
- [3] S. Boughnoux. From projective to euclidean space under any practical situation, a criticism of self-calibration. In *Proceedings of the 6th International Conference on Computer Vision, Bombay, Inde*, pages 790–796, 1998.
- [4] O. Faugeras et F. Lustman. Motion and structure from motion in a piecewise planar environment. *International Journal of Pattern Recognition and Artificial Intelligence*, 2(3):485–508, septembre 1988.
- [5] R. Hartley et P. Sturm. Triangulation. *Computer Vision and Image Understanding*, 68(2):146–157, 1997.
- [6] R.I. Hartley et A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, juin 2000.
- [7] Q.T. Luong. *Matrice fondamentale et autocalibration en vision par ordinateur*. Thèse de doctorat, Université de Paris-Sud, Orsay, décembre 1992.
- [8] Q.T. Luong et T. Vieville. Canonic representations for the geometries of multiple projective views. *Computer Vision and Image Understanding*, 64(2):193–229, 1996.
- [9] J.J. Moré, B.S. Garbow, et K.E. Hillstrom. User guide for *minpack-1*. Rapport technique ANL-80-74, Argonne National Laboratory, USA, 1980.
- [10] P. Sturm. *Vision 3D non calibrée : contributions à la reconstruction projective et étude des mouvements critiques pour l'auto-calibrage*. Thèse de doctorat, Institut National Polytechnique de Grenoble, décembre 1997.
- [11] R. Szeliski et P.H.S. Torr. Geometrically constrained structure from motion : Points on planes. In *3D Structure from Multiple Images of Large-scale Environments SMILE'98*. Springer Verlag, juin 1998.
- [12] J.-P. Tarel et J.-M. Vézien. A generic approach for planar patches stereo reconstruction. In *Proceedings of the Scandinavian Conference on Image Analysis*, pages 1061–1070, Uppsala, Suède, 1995.
- [13] B. Triggs. Optimal estimation of matching constraints. In *3D Structure from Multiple Images of Large-scale Environments SMILE'98*. Springer Verlag, juin 1998.
- [14] G. Xu, J.-I. Terai, et H.-Y. Shum. A linear algorithm for camera self-calibration, motion and structure recovery for multi-planar scenes from two perspective images. In *Proceedings of the Conference on Computer Vision and Pattern Recognition, Hilton Head Island, South Carolina, USA*, juin 2000.