

Méthode de RECONSTRUCTION 3D de Faces Planes par STEREOVISION utilisant les CONTRAINTES de PLANARITE

(3D Reconstruction Method of Plane Faces by Stereovision using the Planarity Constraints)

Patrick Bonnin *, Bertrand Zavidovique **

* ETCA/CREA/SP 16 bis Av Prieur de la Côte d'Or 94114 Arcueil Cédex France
tel (16 1) 42 31 97 59

** ETCA et Institut d'Electronique Fondamentale (IEF) bât 220 Université Paris Sud ORSAY

RÉSUMÉ

Nous présentons une méthode de Reconstruction 3D de facettes planes, basée sur les CONTRAINTES DE PLANARITE. Une nouvelle implémentation de ces contraintes, spécifique à notre méthodologie de segmentation coopérative contours / régions est ici détaillée.

Peu utilisées dans la littérature, car d'essence globale, nous proposons une implémentation mixte globale / locale de ces contraintes.

Les avantages de cette méthode sont au niveau conceptuel sa simplicité, la maîtrise des incertitudes, et au niveau opérationnel sa rapidité et la précision de la reconstruction 3D obtenue.

1. Cadre Applicatif

Le cadre de cette étude se limite à un type de scènes bien précis: scènes d'intérieur (halls, bureaux, etc ...), ou d'extérieurs urbains (composées essentiellement de bâtiments). Les objets présents dans ces scènes sont généralement modélisables d'une manière géométrique simple sous forme de polyèdres (un polyèdre est un volume limité par un ensemble de plans). Ainsi, une méthode simple d'obtention de la modélisation surfacique des polyèdres consiste à reconstruire l'ensemble des plans les constituant. D'où l'intérêt de la méthode que nous proposons, permettant d'obtenir de façon rapide et précise l'équation de facettes planes 3D à partir d'un couple d'images stéréoscopiques.

Nous nous plaçons dans le contexte de la robotique mobile. Notre objectif est d'utiliser la vision à des fins de navigation. Certes, nous ne prétendons pas à la métrologie avec un tel système dans ce contexte, mais nous devons connaître la précision de la reconstruction 3D obtenue, ainsi que les limites du système. Notre démarche va être totalement guidée par des notions de précision et d'incertitude sur les mesures.

Pour ce faire, nous avons monté une première manipulation (d'autres vont suivre), avec un banc stéréo dans un hall. La scène observée est "parfaitement" connue (à la précision des mesures près c'est-à-dire à quelques centimètres). La qualité de la technique est évaluée par comparaison entre

ABSTRACT

A 3D reconstruction method of plane faces, based on the planarity constraint is here described. A new implantation of these constraints, adapted to the use of our edge point / region cooperative segmentation is proposed.

Rarely used on papers, because of their global nature, an mixte global / local implementation of these constraints is proposed.

The advantages of this method are at the conceptual level its simplicity, the control of the uncertainty, and at the operationnal level, its rapidity and the accuracy of the obtained 3D reconstruction.

les données réelles mesurées, et les données reconstruites des objets de la scène.

2. Notions d'Incertitudes

L'incertitude réfère aux instruments de mesure, c'est-à-dire dans notre cas, se manifeste au niveau des images (les instruments de mesure étant évidemment les caméras). Le premier type d'incertitude provient de la discrétisation du plan focal des caméras (traduction de coordonnées continues en nombre de pixels). Le second, provient de la segmentation: les primitives de l'image doivent être précisément localisées. Pour ce faire, dans notre méthode de segmentation coopérative mixte contours / régions [BZ90], l'information issue de la détection de contour est privilégiée en ce qui concerne la localisation. Le troisième et dernier type, venant s'ajouter aux deux précédents provient de la précision des réglages (caméras rigoureusement parallèles, dans un plan horizontal, et perpendiculaires à leur base) et la connaissance des paramètres des caméras et du banc stéréo.

Nous supposerons l'incertitude résultante isotrope en projection sur l'image, et pouvant être estimés par un nombre de pixels δn . Reportons cette incertitude dans le plan 3D (cf figure n°1).

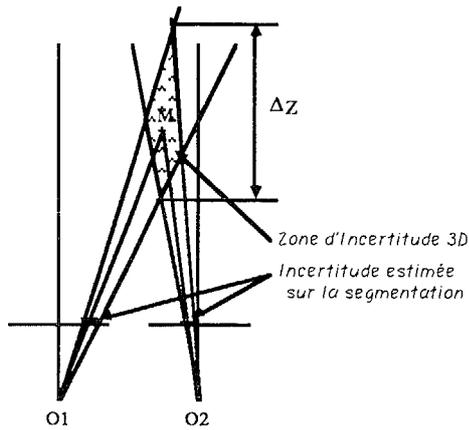


Figure n°1: Calcul de l'incertitude 3D

Par exemple sur l'axe OZ:

$$Z_{\max} = \Delta f / [\delta x (dx - 2 \delta n)]$$

$$Z_{\min} = \Delta f / [\delta x (dx + 2 \delta n)]$$

- avec
- f: focale des caméras,
 - Δ: distance entre les caméras,
 - δx, δy : taille en x et y du pixel sur le plan focal
 - ngx et ndx coordonnées pixel en X sur les images gauche et droite
 - dx = ngx - ndx disparité

d'où	$\Delta Z = Z_{\max} - Z_{\min} = \frac{AZ^2}{(1 - BZ^2)}$
avec	$A = 4 \delta n \delta x / f \Delta$
et	$B = [2 \delta n \delta x / f \Delta]^2$

Compte tenu des ordres de grandeur: $B \ll 1$, ainsi pour des conditions usuelles l'erreur ΔZ est quasiment proportionnelle au carré de la distance du point M de l'espace 3D aux plan des caméras. En d'autres termes, ΔZ n'est pas constante en fonction de l'endroit où se trouve le point observé. Un développement analogue montre qu'il en est de même pour ΔX et ΔY . De plus ΔX , ΔY et ΔZ sont différents, ce qui veut dire que l'incertitude de localisation autour d'un point 3D n'est absolument pas isotrope.

Or, l'hypothèse d'une incertitude isotrope et constante en fonction du lieu du point 3D considérée est à la base de méthodes locales de reconstruction 3D où le plan est obtenu par minimisation de sa distance à un ensemble de points 3D, obtenus par triangulation à partir des primitives locales (segments ou jonctions de contours) des deux images. Néanmoins, ces méthodes peuvent suffire dans le cas de facettes planes de tailles relativement peu importantes, et quasiment parallèles au plan des caméras. Dans un cas contraire, cela reviendrait à dire que l'on accorde autant de crédibilité (pour ce qui est de la précision) à un point très proche par exemple situé à 1m de distance qu'à un point plus lointain, par exemple situé à 50 m.

La seule manière d'obtenir l'équation de la facette plane, par une méthode d'optimisation rigoureuse, tout maîtrisant l'incertitude, est de rester au niveau 2D des images.

3. Détermination de l'Equation de la Facette Plane Directement à partir des Images

Dans [Gag89], pour un couple de régions associées, et pour les projections d'un plan d'équation $Z = aX + bY + c$, il est possible de prédire la position de la région droite R_{dg} (respectivement gauche R_{gd}) à partir des paramètres de l'équation du plan et de la région gauche R_g (resp droite R_d). Le choix des paramètres optimaux de l'équation du plan a^*, b^*, c^* , est une minimisation de l'erreur quadratique entre les vecteurs d'attributs des régions et de leurs prédictions (aire, centre de gravité, axes principaux d'inertie). La minimisation de

$\| \text{Car R}_{dg} - \text{Car R}_d \|^2 + \| \text{Car R}_{gd} - \text{Car R}_g \|^2$ est obtenue par une technique de relaxation. La contrainte de planarité est ici implicite.

[RML87] utilise explicitement la relation que nous appellerons la **contrainte de planarité**. Celle-ci assure l'existence d'une transformation affine entre les coordonnées de points de l'image gauche (X_g, Y_g) et de points de l'image droite correspondants (X_d, Y_d) provenant de points 3D de la scène (X, Y, Z) situés sur une même surface plane. En d'autres termes, pour un banc stéréo à géométrie rectifiée, la contrainte s'exprime simplement par:

$X_d = \alpha X_g + \beta Y_g + \gamma$	$Y_d = Y_g$
---	-------------

Les coefficients α, β, γ dépendent des paramètres a, b, c, d de l'équation du plan contenant la surface plane $aX + bY + cZ + d = 0$. Ces coefficients induisent des relations entre les attributs (aire, centre de gravité, moments d'inertie) des régions du couple associées entre les deux images. Dans cet même article, les coefficients α, β, γ sont obtenus par minimisation d'un critère d'erreur quadratique entre les vecteurs d'attributs de la région droite et de sa prédiction à partir des coefficients α, β, γ et de la région gauche. La cohérence est vérifiée en comparant (par superposition géométrique et photométrique) la région droite et sa prédiction.

La méthode que nous proposons est une autre mise en œuvre de cette contrainte, mettant à profit la richesse de l'information extraite par la segmentation coopérative points de contours / régions. Outre les coefficients a, b, c, d de l'équation du plan contenant la facette plane, cette nouvelle implantation permet d'obtenir en plus les sommets 3D de cette facette, ce qui n'est pas le cas dans la méthode précédente.

4. Mise en Œuvre

4.1. Implémentation Mixte Globale / Locale

Les contraintes de Planarité sont des contraintes d'essence globale, puisqu'elle assurent l'existence d'une et une seule transformation affine (ie. coefficients α, β, γ) pour tous les points d'une même facette 3D. Néanmoins, une implémentation mixte globale / locale est possible.

En effet, supposons que l'on dispose d'un ensemble de couples de points $\{(m_{ig}, m_{id}) / \forall i \in [1, n]\}$. De plus, supposons que chaque couple est formé d'un point de l'image gauche m_{ig} et droite m_{id} correspondant à un même point M_i de l'espace 3D. Il est ainsi possible de déterminer les coefficients α, β, γ par minimisation dans l'image droite de l'erreur moyenne entre l'abscisse du point x_{id} de m_{id} et son

estimation à partir de m_{ig} par la contrainte de planarité. Le triplet $\alpha^*, \beta^*, \gamma^*$ retenu est celui qui minimise l'expression:

$$\varepsilon^2 = \sum_{i=1}^n (x_{id} - \alpha x_{ig} - \beta y_{ig} - \gamma)^2$$

En fait nous disposons de cet ensemble de couples de points. En effet, la structure de données issue de l'algorithme de segmentation coopérative mixte régions / contours, est un graphe d'adjacence de régions, dont la particularité est de stocker également les points de jonctions de segments propres à chaque région. La méthode de mise en correspondance entre les graphes des deux images est hiérarchique et mobilise deux niveaux. Le haut est constitué des primitives *globales* (ie les régions) de l'image; le bas, des primitives *locales* (ie les points de jonctions de segments). L'association à haut niveau contraint l'association à bas niveau, en le sens que seuls les points de jonctions appartenant au couple de régions associées à haut niveau seront pris en compte pour l'association locale.

4.2. Prise en Compte du Contexte de la Robotique Mobile

Compte tenu de la probabilité de présence des divers types de plans dans une scène de "robotique mobile" en environnement intérieur ou extérieur urbain, le système recherche en priorité les plans verticaux, puis les plans horizontaux, puis en dernier les plans obliques. La mobilité conduit à distinguer également plusieurs types parmi les plans verticaux. Ainsi le système recherchera les plans perpendiculaires au déplacement, puis les plans parallèles, puis les plans en travers, ce pour obtenir une meilleure et surtout plus efficace connaissance de l'environnement. Un plan parallèle au déplacement sera moins gênant pour une translation du robot qu'un plan orthogonal.

4.3. Expression des Diverses Hypothèses

Compte tenu des repères choisis (cf figure n°2), les relations entre a,b,c,d et α, β, γ sont:

$a = (\alpha - 1) / \Delta$	$b = \beta / \Delta$
$c = \gamma / (f \Delta)$	$d = 1 + (\alpha - 1) / 2 + h\beta / \Delta$

Les calculs seront détaillés dans [Bon91]

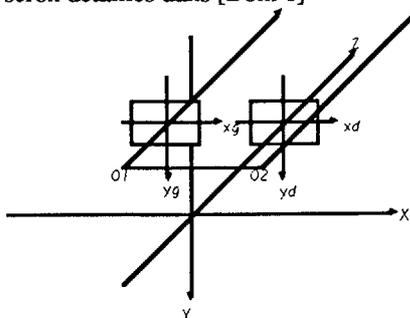


figure n°2 Banc Stéréo et Repères utilisés

Pour chaque hypothèse les formules permettant de déterminer les coefficients α, β, γ sont:

* Plan vertical parallèle à XOY (Z = cte)

$a = 0, b = 0$ ce qui implique $\alpha = 1, \beta = 0,$
 $\gamma = (Xd - Xg) / n$ avec n nombre de

points de jonction associés

$$X_d = \sum_{i=1}^n x_{id} \quad \text{et} \quad X_g = \sum_{i=1}^n x_{ig}$$

* Plan Vertical parallèle à YOZ (X = cte)

$c = 0, b = 0$ ce qui implique $\beta = 0, \gamma = 0,$

$\alpha = X_{gd}^2 / X_{gg}^2$ avec

$$X_{gd}^2 = \sum_{i=1}^n x_{ig} x_{id} \quad X_{gg}^2 = \sum_{i=1}^n x_{ig} x_{ig}$$

* Plan Vertical Quelconque

$b = 0,$ ce qui implique $\beta = 0,$

$\alpha = (nX_{gd}^2 - X_d X_g) / (nX_{gg}^2 - (X_g)^2)$

$\gamma = (X_{gg}^2 X_d - X_{gd}^2 X_g) / (nX_{gg}^2 - (X_g)^2)$

* Plan Horizontal

$a = 0, c = 0,$ ce qui implique $\alpha = 1$ et $\gamma = 0,$

$\beta = (YX_{gd} - YX_{gg}) / Y_{gg}^2$ avec

$$YX_{gd} = \sum_{i=1}^n y_{ig} x_{id} \quad YX_{gg} = \sum_{i=1}^n y_{ig} x_{ig} \quad Y_{gg}^2 = \sum_{i=1}^n y_{ig} y_{ig}$$

* Plan quelconque

Aucune restriction sur les paramètres a,b,c et d, d'où l'on détermine α, β, γ

$$\alpha = [(X_{gd}^2 - X_g X_d) (Y_{gg}^2 - (Y_g)^2) - (X_{Ygd} - X_d Y_g / n) (X_{Ygg} - Y_g X_g / n)] / [(X_{gg}^2 - (X_g)^2 / n) (Y_{gg}^2 - (Y_g)^2 / n) - (YX_{gg} - X_g Y_g / n)^2]$$

$$\beta = [(YX_{gd} - Y_g X_d / n) (X_{gg}^2 - (X_g)^2 / n) - (X_{gd}^2 - X_d X_g / n) (X_{Ygg} - Y_g X_g / n)] / [(X_{gg}^2 - (X_g)^2 / n) (Y_{gg}^2 - (Y_g)^2 / n) - (YX_{gg} - X_g Y_g / n)^2]$$

$\gamma = (X_d - \alpha X_g - \beta Y_g) / n$

4.4. Vérifications des Hypothèses

La vérification de la cohérence de la facette reconstruite s'effectue en deux étapes. La première, dans l'image droite est très lapidaire. Elle permet un premier rejet, très rapide des fausses hypothèses. Elle consiste à vérifier si l'écart entre X_d , coordonnée selon X dans l'image droite et X_{dp} sa valeur prédite, est satisfaisant, ceci pour chaque couple de points associés:

$$|X_d - \alpha X_g - \beta Y_g - \gamma| < \delta n$$

La seconde, dans le plan 3D est plus longue. Elle consiste, en un premier temps à reconstruire le point 3D M, dont les projections M_g et M_d ont été associées dans les images, puis à calculer autour de chaque point M l'espace d'incertitudes sur les coordonnées 3D X,Y,Z, connaissant l'incertitude δn au niveau de l'image. Ces deux étapes, totalement indépendantes de l'hypothèse générée, ne sont accomplies qu'une fois au début de l'algorithme. La troisième, spécifique à chaque hypothèse consiste à vérifier, pour chaque point M, si sa projection sur la facette plane générée appartient à son espace d'incertitude. Dans la négative, l'hypothèse est rejetée.



4.5. Possibilité de Gestion de ces Hypothèses

Pour une facette plane suffisamment significative (ie taille en fonction de la distance, ou mieux angle solide de vue), les mauvaises hypothèses sont rejetées violemment (coûts très importants). Une stratégie consistant à générer hypothèse après hypothèse tout en vérifiant, dans l'ordre donné ci-dessous, s'arrêtant dès qu'une hypothèse est validée, semblerait tout à fait satisfaisante. Notons le double avantage d'une telle stratégie:

- rapidité, car les premières hypothèses se trouvent être les plus simples à évaluer
- obtention immédiate de connaissances relatives à la facette plane observée: la classification est directe.

5. Résultats

Un schéma de la scène traitée est représenté figure n° 3. Elle est constituée de deux petits meubles, d'un plan vertical oblique, et d'une cloison de séparation. A gauche sont données les informations de distance.

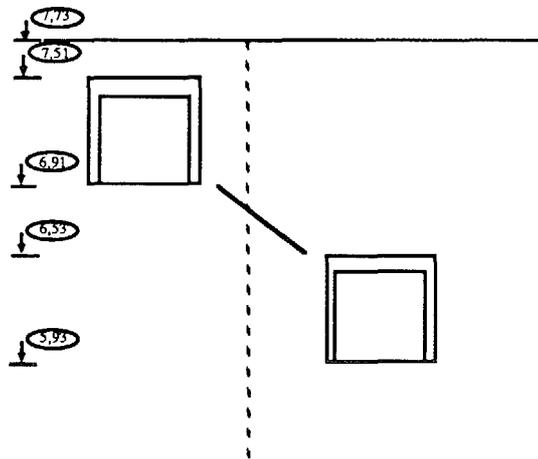


figure n° 3: Scène traitée

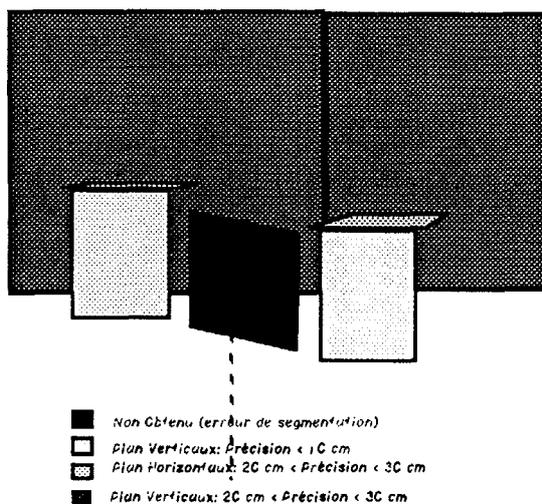


figure n°4: Résultats obtenus

Les projections du plan oblique dans les images n'ont pu être extraites par la phase de segmentation. La raison en est qu'il n'y a pas une transition assez franche avec le sol. Par conséquent, ce plan n'a pas pu être correctement reconstruit.

En revanche, pour les plans verticaux orthogonaux (meubles et cloison de séparation), la précision est très correcte (cf figure n° 4) pour l'utilisation que l'on envisage d'un tel système. Cette précision est inférieure à 10 cm pour les plans verticaux des meubles situés à des distances de 6m et 7m. Elle est comprise entre 20cm et 30cm pour les plans de séparation situés à 8m. Il faudrait de plus nombreuses mesures pour vérifier expérimentalement si l'incertitude sur Z suit la loi "théorique" énoncée précédemment $\Delta z = AZ^2 / (1 - BZ^2)$, qui croit plus vite en fonction de Z que la fonction AZ^2 . Notons que la vérification dans l'espace 2D est très sélective: son coût est de l'ordre du pixel pour les hypothèses correctes, et supérieur à 10 pixel pour les hypothèses erronées (voir 40 pour certaines!).

6. Conclusion

Nous avons exposé en détails une méthode de reconstruction 3D de facettes planes, à partir d'un couple d'images stéréoscopiques. Celle-ci a de nombreux avantages:

- simplicité: les calculs peuvent être menés au bout "à la main",
- rapidité: de par la stratégie de gestion des hypothèses (le premier rejet est très rapide et les hypothèses sont générées par l'ordre décroissant des probabilités d'apparition dans la scène: plan vertical, horizontal puis oblique),
- précision: par la prise en compte du fait que l'incertitude autour d'un point 3D dépend de la localisation de celui-ci, et que par conséquent, il fallait utiliser une autre méthode, indépendante de la position du point dans l'espace, ce qui ne peut être qu'en travaillant au niveau de l'image
- classification des facettes planes: déterminante pour la reconstruction des polyèdres.

La suite de cette étude consistera d'une part à reconstruire effectivement les polyèdres 3D à partir des facettes planes ainsi constituées, et d'autre part, à trouver une stratégie de retour sur les hypothèses d'associations entre points des images lorsqu'aucun type de facette plane n'a pu être correctement reconstruit.

7. Références

- [Bon91] P. Bonnin, thèse de Doctorat à paraître en fin 91
- [BZ90] P. Bonnin, B. Zavidovique
Segmentation coopérative mixte région / contour spécifique à la reconstruction 3D de scènes
Journées d'étude TDF/SEE "la télévision en relief" 25/26 avril 1990 CCETT Rennes
- [Gag89] A. Gagalowicz (conférencié invité)
Coopération entre l'analyse et la synthèse d'images
7ième Congrès RFIA Paris Nov 1989
- [RML87] Y. Remion, H. Maître, J. Lopez Krahe
Recalage de zones de 2 vues stéréoscopiques d'un univers composé d'objets plans
Congrès COGNITIVA 87 18 / 22 mai 1987 pp 174-180

8. Remerciements

Cette étude a été financée par le groupe n°1 du Service des Recherches de la Direction des Recherches et Etudes Techniques au titre du contrat DRET/SDR/G1 n° 89/1027.