



# Reconstruction tridimensionnelle des bâtiments par une méthode de stéréovision fondée sur la géométrie projective

M. ZAHID, M. CHAPRON, J.P. COCQUEREZ

ETIS-ENSEA, Allée des chênes pourpres, 95014 Cergy Pontoise

Tel. : (1) 30 73 66 12 Fax : (1) 30 73 66 27

## RESUME

Dans cet article, nous exposons les différentes étapes permettant à un système d'analyse d'images de reconstruire en 3D des bâtiments à partir d'une paire d'images stéréoscopiques. Après la modélisation géométrique des deux prises de vues, on identifie chaque objet par les projections de ses sommets sur une image, l'image OUEST par exemple. Ils représentent les intersections des portions de contours considérées comme des segments de droite, leurs homologues dans l'image EST étant calculés par une transformation affine qui est une approximation d'une homographie. Une fois l'appariement effectué, on détermine les positions 3-D des points appariés ainsi que les plans correspondant aux différentes facettes des objets.

## ABSTRACT

In this paper, we present the different steps which permits to an image analysis system to reconstruct buildings from a pair of two stereoscopic images a 3-D image. After the geometrical modelling, each object is identified by its vertice projections on one of the two images (west image for example). They represent intersections of edge portions considered as line segments. Their homologous segments are computed by an affine transformation which approximates the homography transformation. Once the matching is performed, 3-D position of matched points are determined. Finally, we determine in the same way the corresponding planes of object faces.

## 1 Introduction

Le but de ce travail est de reconstruire une série de bâtiments à partir d'une paire d'images stéréoscopiques. Les images sont prises par une caméra placée sur un avion : l'avion effectue la première prise de vue, puis il se déplace dans une direction horizontale pour enregistrer la deuxième image de sorte que 60 % de la scène visée sera commune aux deux images. Les bâtiments sont fournis par un système d'analyse d'images ; ils sont représentés sur une seule image (image OUEST) par des régions localisant leurs toits.

En imagerie aérienne et satellitaire, la plupart des méthodes de reconstruction effectuent la modélisation géométrique de la caméra, puis elles déterminent les primitives homologues en utilisant une mise en correspondance par corrélation sur les images épipolaires, ensuite elles reconstruisent les points objets à partir de leurs projections sur les deux plans-images ([Wu 89] et [Maître 88]). Les résultats sont satisfaisants pour les images de zones rurales ou montagneuses, car la corrélation suppose que les surfaces physiques sont localement continues et que les images sont suffisamment texturées. Or les images de bâtiments contiennent des ruptures et les objets sont caractérisés par des plages de niveaux de gris uniformes. L'idée la plus naturelle est de s'orienter vers les approches utilisées en vision robotique fondées sur la mise en correspondance des primitives de type régions ([Mohan 89] et [Lee 89]), segments de droite [Ayache 87] ou jonctions [Herman 86]. Il y a aussi les travaux de Bonnin

et Zavidovique [Bonnin 91] qui utilisent la contrainte de coplanarité pour déterminer les facettes planes des objets en utilisant la prédiction et la vérification d'hypothèses. Ces techniques donnent de bons résultats si la segmentation des deux images est identique, c'est-à-dire, si les deux images sont enregistrées dans les mêmes conditions météorologiques. Du fait que les images ne sont pas enregistrées dans les mêmes conditions et qu'elles contiennent des zones de discontinuités, nous avons utilisé la corrélation et certaines propriétés de la géométrie projective pour effectuer la mise en correspondance.

Dans un premier temps, nous décrivons la modélisation géométrique de la caméra et la reconstruction d'un point objet, puis nous présentons la démarche d'extraction des sommets de bâtiments. Nous exposons ensuite la technique employée pour effectuer la mise en correspondance, et finalement la reconstruction d'un bâtiment ainsi que quelques résultats expérimentaux.

Cette approche a été testée sur des images fournies par l'IGN. Elles sont prises à une hauteur de 4,5 Km environ par une caméra de distance focale 213,28 mm. La résolution de l'image est de l'ordre du mètre.

## 2 Modélisation géométrique de la caméra

Pour une image prise par une caméra, les points 3-D de la scène sont projetés dans le plan image. La modélisation géométrique d'une caméra consiste à trouver les coefficients de la matrice de la transformation perspective directe qui relie les



coordonnées d'un point de la scène aux coordonnées de sa projection dans l'image. Cette matrice est souvent appelée matrice de calibrage de la caméra.

### 2-1 Estimation de la matrice de calibrage

Les travaux effectués dans ce domaine peuvent se classer principalement selon deux classes de méthodes. Elles sont fondées, toutes les deux, sur la connaissance des coordonnées d'un ensemble de points objets dans le repère lié à la scène et de celles de leurs projections dans le plan-image.

La première classe de méthodes, dites méthodes des orientations ([Hout 61] et [Bénard 83]). Ce type de méthodes suppose connues la distance focale de la caméra et les coordonnées de la projection du centre optique. Elle est souvent effectuée en deux phases. Pour la première phase, il s'agit de calculer le déplacement rigide permettant de passer du repère de l'image OUEST au repère de l'image EST ; cette transformation est appelée l'orientation relative. La deuxième phase est la phase de orientation absolue ; elle consiste à déterminer la transformation permettant de passer du repère de la scène au repère de l'image.

La deuxième classe de méthodes est souvent utilisée en vision robotique. Elle consiste à calculer les transformations de passage entre le plan-image et la scène pour chacune des caméras [Faugeras 88]. Nous avons choisi cette approche en s'appuyant sur des travaux de Toscani [Toscani 87] et de ceux de Ben Rhouma [Ben Rhouma 84]. La qualité de l'estimation de cette matrice dépend de la position des points d'appui, du calcul matriciel et de la présence des données aberrantes dans l'ensemble des points d'appui.

En ce qui concerne la position des points d'appui, nous les avons choisis de sorte que les points objets soient sur le bord de la scène. Quant au calcul matriciel, nous avons utilisé la méthode des transformations orthogonales pour la résolution des systèmes linéaires. Pour minimiser l'influence des données aberrantes, nous avons employé un procédé itératif qui met en oeuvre la technique des moindres carrés pondérés. Ceci a permis d'améliorer la précision de reconstruction de 50%.

Notons qu'il y a des déformations qui peuvent affecter l'objectif de la caméra. Parmi l'ensemble des déformations affectant les lentilles de l'appareil, on considère seulement les erreurs de montage de la caméra et la distorsion géométrique [Slama 80] qui entraînent une modification des positions théoriques des points images. Ces modifications diminuent avec la distance qui sépare la caméra de la scène, or dans les images aériennes cette distance est très importante ; on peut donc négliger ces déformations. Les autres déformations ont des effets optiques, elles affectent donc seulement la qualité de l'image.

### 2-2 Reconstruction tridimensionnelle d'un point objet

Pour retrouver les coordonnées d'un point de la scène, il faut disposer de ses deux projections dans les deux plans-

images. Dans ce cas, le point cherché sera l'intersection des deux rayons de projection. Du fait que ces rayons sont déterminés à partir des coefficients des 2 matrices de calibrage, leurs paramètres sont imprécis : donc il peut arriver que deux rayons de projection issus du même point de la scène ne soient pas sécants. Dans ce cas, on choisit le point le plus proche de ces deux rayons ; il correspond au milieu leur perpendiculaire commune. Les coordonnées de ce point sont la solution d'un système surdéterminé dont les équations sont celles des deux rayons de projection.

### 3 Extraction des sommets d'un bâtiment de l'image OUEST

Pour extraire les sommets d'une région représentant un bâtiment de l'image OUEST, on identifie tout d'abord les parties du contour susceptibles d'être des segments de droite en utilisant la technique de Bresenham [Bresenham 65]. Nous déterminons ensuite les sommets par concaténation des segments détectés.

#### 4 Mise en correspondance

La stéréovision est un moyen de recueillir l'information 3-D d'une scène par application des règles géométriques qui nécessitent une mise en correspondance préalable des primitives. La mise en correspondance consiste à identifier des primitives homologues issues d'un même objet. Plusieurs algorithmes ont été proposés, ils diffèrent sur trois points fondamentaux : les contraintes utilisées pour réduire le temps de calcul, le choix des primitives à mettre en correspondance et la démarche suivie pour apparier les primitives choisies [Pollard 85]. On trouvera dans [Baker 82] une étude critique des principales méthodes développées dans les années 1970. Compte tenu de l'instabilité des résultats de la segmentation par le traitement bas-niveau des images et de la discontinuité des zones où se situent les points dont on veut déterminer les homologues (les sommets de bâtiments), on utilise le théorème fondamental de la géométrie projective [Berger 77] qui permet de déterminer une fonction  $f_p$  du plan-image OUEST vers le plan-image EST. Cette application associe à chaque point du plan-image OUEST son homologue dans le plan-image EST. C'est une homographie que l'on peut déterminer à partir de quatre points formant un quadrilatère dans l'image OUEST et de leurs homologues dans l'image EST [Zahid 92]. Celle-ci est alors utilisée pour déterminer les homologues des sommets.

#### 4-1 Sélection des points de repères

L'évaluation de l'homographie  $f_p$  est d'autant plus précise qu'on est proche de l'un des quatre points. Elle est également d'autant plus précise que les points utilisés sont éloignés les uns des autres. L'idée la plus naturelle est donc d'utiliser les points situés aux coins des deux images. Mais les erreurs sur les points utilisés pour déterminer les coefficients de sa matrice sont de l'ordre d'un pixel ; ces erreurs sont énormes pour le calcul matriciel permettant cette identification. Pour réduire leur

influence, nous avons déterminé une fonction pour chacune des régions correspondant à un toit de bâtiment, car cette homographie donne de bons résultats seulement pour les points situés dans un voisinage des points utilisés pour le calcul de la matrice. Pour cela, on choisit un ensemble de points dont représente le centre de la partie du toit incluse dans une fenêtre centrée sur l'un des sommets du bâtiment à reconstruire. Ils sont situés dans des zones continues dont la variation des niveaux de gris est importante. Par conséquent, nous pouvons déterminer leurs homologues par corrélation.

#### 4-2 Recherche des homologues des points de repère

On peut déterminer ces homologues par corrélation en utilisant la contrainte d'épipolarité. La fonction de corrélation utilisée est une fonction qui associe le plus grand nombre au couple de points issus d'un même point objet, elle est centrée et normalisée [Wu 89]. Du fait qu'on applique la corrélation pour des points voisins de zones de discontinuités, il peut arriver que la fenêtre de corrélation soit à cheval sur deux surfaces de hauteurs différentes. Pour éviter ce problème, on prend en compte seulement la partie de la fenêtre située dans la région dont on veut déterminer les homologues des sommets. Pour cela, on multiplie les termes de la fonction de corrélation par une application indicatrice  $1_{\mathbb{R}}$ . L'application  $1_{\mathbb{R}}$  associe aux points intérieurs la valeur 1 et aux autres points 0.

#### 4-3 Estimation d'une approximation de la transformation fp

Le déplacement rigide permettant de passer de l'image OUEST à l'image EST est presque une translation horizontale (Orientation relative [Cocquerez 90]). Donc, le système de prise de vue en imagerie aérienne est tel que la forme et les attributs des toits de bâtiments sont invariants par la transformation perspective. En conséquence, l'homographie fp peut être approximée par une transformation affine Tr. Au lieu de déterminer l'automorphisme associé à l'homographie fp, ce qui demande beaucoup de calculs, on détermine directement la transformation Tr. Afin d'éviter le calcul matriciel qui pose le problème de la stabilité numérique, on détermine les paramètres de l'application Tr par la règle de Widrow et Hoff [Widrow 60].

Il peut arriver que le maximum de la fonction de corrélation f soit atteint en deux points p et p' qui ne sont pas homologues. Par conséquent, les homologues trouvés par la corrélation et les homologues déterminés par l'application Tr ne coïncident pas toujours. Pour avoir une bonne précision, on procède de la façon suivante : une fois calculée la transformation entre les deux régions, on compare les homologues trouvés par la corrélation avec les points calculés à partir de ceux de l'image OUEST par l'application Tr ; s'ils ne sont pas voisins, on effectue la recherche de ces homologues avec une fenêtre de taille plus grande, on calcule alors de nouveau la transformation Tr, puis on

itére le processus jusqu'à ce que la précision ne s'améliore plus ou que la taille de la fenêtre devienne trop importante.

#### 4-4 Seuils

La mise en correspondance met en oeuvre deux seuils qui sont calculés de manière automatique. Le premier seuil concerne la largeur de la bande autour de la droite épipolaire où s'effectue la recherche de l'homologue dans l'image EST d'un sommet de l'image OUEST. Quant au deuxième seuil, c'est la taille de la fenêtre de la corrélation.

### 5 Reconstruction des bâtiments

Nous avons supposé que les bâtiments ont la forme d'un prisme droit, donc on peut le construire à partir des sommets de son toit. Une fois les sommets des bâtiments à reconstruire sont identifiés dans les deux images, on détermine les points objets correspondant. Nous calculons ensuite le plan représentant le toit du bâtiment ; ce plan est le plus proche de tous les sommets, puis nous reconstruisons le bâtiment à partir des projections orthogonales des sommets retrouvés par le calcul sur le plan représentant le toit du bâtiment, puis on reproduit l'image des bâtiments vues dans une direction quelconque à partir d'un endroit quelconque.

Pour retrouver les coordonnées 3-D d'un point objet nous avons pris deux précautions : la première concerne la minimisation des erreurs de troncature, la deuxième consiste à n'utiliser que les équations qui entraînent le minimum d'erreurs.

### 6 Résultats expérimentaux

Les photos 1 et 2 illustrent la méthode de mise en correspondance sur un bâtiment. La photo 3 montre un exemple de reconstruction de bâtiments.

### 7 Conclusion

Ce travail apporte 2 contributions essentielles :

- la première réside dans l'amélioration de la précision de la modélisation géométrique grâce à la technique des moindres carrés pondérés, au choix des points d'appui situés sur le bord de la scène et à l'utilisation des transformations orthogonales.
- la deuxième consiste à utiliser la géométrie projective qui permet de déterminer une transformation affine entre les deux plans-images. Celle-ci a permis la détermination des homologues des sommets avec une bonne précision, l'identification des faux appariements, et le calcul des coordonnées des points cachés, ou des points que la corrélation n'a pas pu trouver, ce qui est souvent le cas dans les images aériennes de zones péri-urbaines.

La précision de reconstruction est voisine de celle des points d'appui, donc elle ne peut être améliorée. Cette approche est complémentaire des méthodes classiques de reconstruction de modèles numériques de terrain : on reconstruit la scène sans les bâtiments avec les méthodes existantes, puis on reconstruit les bâtiments par notre approche dans le but d'obtenir une meilleure précision.



### Références

- [Ayache 87] N. Ayache, O. D. Lustman, Fast and reliable passive stereovision using three cameras, International Workshop on Machine Vision and Machine Intelligence, Japan, Febr. 1987.
- [Baker 82] H. H. Baker, Depth from Edge and intensity based stereo, Stanford university, Depart. of Com. Science, Sep. 1983.
- [Ben Rhouma 84] Ben Rhouma, Analyse de scène par une méthode K2D en robotique avancée, Thèse de doctorat en sciences, Paris 11, 1984.
- [Berger 77] M. Berger, Géométrie. Action de groupe, espaces affines et projectifs, CEDIC, Nathan, 1977.
- [Bonnin 91] P. Bonnin, B. Zavidovique, Méthode de reconstruction 3-D de faces planes par stéréovision utilisant la contrainte de coplanarité, 13<sup>ème</sup> GREDSI, Juan-les-Pins, Sept. 1991.
- [Bresenham 65] J-E. Bresenham, Algorithm for computer control of a digital plotter, IBM System Journal, Vol. 4, p. 25-30, 1965.
- [Cocquerez 90] J.P. Cocquerez, M. Jordan, M. Zahid, Geometric Modelling of Stereovision : Application to aerial pictures, 9<sup>th</sup> ECAI, Stockholm- Sweden, Aug. 1990.
- [Hout 61] O. M. A. Van Den Hout, Analytical orientation methods, Bolletino de geodesia scienze, Affini, Vol. xx 3, 1961.
- [Faugeras 88] O. D. Faugeras, Quelques pas vers la vision artificielle en trois dimensions, Tech. Scen. Infor., Vol. 7, p. 547-690, 1988.
- [Herman 86] M. Herman and T. Kanade, Incremental reconstruction 3-D scene from multiple, complex images, AI, No 30, p. 333 - 341, 1986.
- [Larabi 91] S. Larabi, B. Thiesse, P. Marthon, Utilisation de la géométrie projective en vision stéréoscopique binoculaire : Appariement de stéréogrammes de segments, 8<sup>ème</sup> AFCET 91, Lyon-France, 25-29 Novembre 1991.
- [Lee 90] H. J. Lee and W. L. Lei, Region matching and depth finding for 3-D objects in stereo aerial photographs, Pattern Recognition, Vol. 23, p. 81-94, 1990.
- [Maître 88] H. Maître, C. Bardinnet, "Digital stereo relief reconstruction to Tanzania using arial imagery, Image Anal. Geol. Con. Radi Surv., Musée Royal de l'Afrique Central, N° 96, 12, 1988.
- [Mohan 89] R. Mohan, R. Nevatia, Using perceptual organization to extract 3-D structures, Pat. Anal. and Mach. Int., Vol. 11, p. 1121- 1139, 1989.
- [Morin 91] L. Morin, R. Mohr E. Grosso, Positionnement relatif à partir d'invariants projectifs, 8<sup>ème</sup> AFCET 91, Lyon-France, 25-29 Novembre 1991.
- [Pollard 85] S. Pollard, J. E. W. Mayhen, J. P. Frisby, PMF : A Stereo correpondance algorithm using a disparity gradient limit, Perception, Vol. 14, p. 449-470, 1985.
- [Slama 80] C. Slama, Manual of photogrammetry, American society of photogrammetry, falls church edition, 1980.
- [Widrow 60] G. Widrow, M. E. Houff, Adaptation of radio Engineers, Western, Electronic show and convention Record, Part 4, p. 96-104, 1960.
- [Toscani 87] G. Toscani, Système de calibration et perception du mouvement en vision artificielle, Thèse de doctorat en sciences, Paris 11, 1987.
- [Wu 89] Y. Wu, H. Maître, Mise en correspondance d'images stéréo par programmation dynamique utilisant la cohérence interligne, 12<sup>ème</sup> GRETSI, Juan-Les-Pins-France, Juin 1989.
- [Zahid 92] M. Zahid, Stéréovision en imagerie aérienne : reconstruction tridimensionnelle des bâtiments, Thèse de doctorat en science de l'Université Paris Sud (centre d'Orsay), Mai 1992. AFCET, Lyon, Nov. 1991.

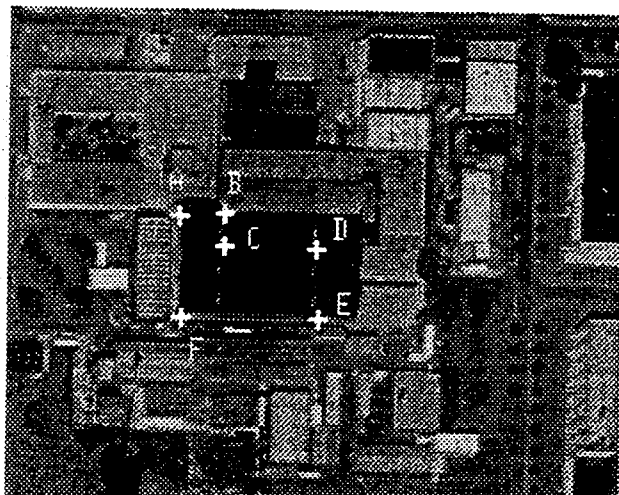


Photo 1

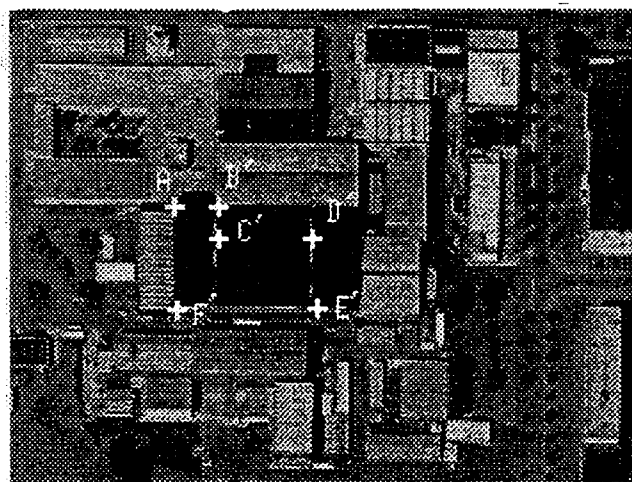


Photo 2 Mise en correspondance avec l'image précédente

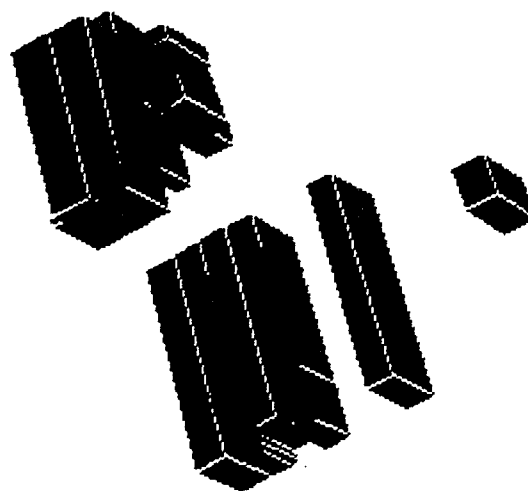


Photo 3 Un exemple de reconstruction 3D