

## STEREOVISION MONOCULAIRE LATÉRALE

R. SAMY, D. DUCLOS, JL. ERNANDEZ

SAT - DOD

41 Rue Cantagrel - BP 389

75626 PARIS CEDEX 13 - FRANCE

RESUME

La stéréovision monoculaire latérale présente de nombreux avantages dans des applications civiles et militaires comparées à la stéréovision axiale pratiquée classiquement en robotique en raison de l'observabilité de la distance.

Cet article présente une évaluation d'une chaîne de reconstruction de bâtiments dans une scène aérienne.

Cette chaîne de traitement comporte une extraction de contours ainsi des estimations robustes du mouvement et de disparités par un filtre de Kalman.

The side-looking monocular stereovision present many advantages in both aerospace and civilian applications compared to classical frontal vision usually explored in vision for robots despite of problems of observability of the range.

This paper describe an evaluation of a processing scheme for the reconstruction of buildings in aerial scene.

This processing scheme include a robust edge detection and also an robust motion estimation and a kalman filter for disparity refinement.

1 - INTRODUCTION

Un des objectifs de la stéréovision par ordinateur est de reconstruire la structure tridimensionnelle (3D) des objets d'une scène à partir d'une ou plusieurs images. La vision stéréoscopique utilise deux images prises avec deux caméras. Connaissant le modèle de projection de chaque caméra et la relation spatiale entre les deux caméras, il s'agit de calculer les coordonnées 3D d'un point à partir de ses deux projections dans les deux images.

Le problème d'appariement ou de mise en correspondance consiste à trouver un ensemble de couples ou de paires, chaque couple étant formé d'un point d'une image apparié avec un point de l'autre image. Ce problème a une nature combinatoire et la seule façon de réduire cette combinatoire est de définir une mesure de ressemblance entre un point d'une image et un point de l'autre image et de mettre en oeuvre des contraintes qui ont trait à la géométrie du capteur stéréoscopique et à la structure de la scène et des objets observés.

Le problème de la reconstruction d'une cible acquise dans une séquence d'images peut être traité de multiples façons.

En vue d'une reconstruction 3D d'objets de petite taille, on s'intéresse dans l'étude présentée aux techniques de stéréovision dans notre contexte d'application sachant que les techniques classiques ne sont pas exploitables.

Les performances des algorithmes de reconstruction 3D des objets de petite taille dépendent beaucoup de la précision de la navigation et de plus doivent être robustes aux changements d'aspects des objets à reconstruire, ceci est essentiellement pris en compte par les contours actifs (voir article GRETSI 95 de JF. Bonnet et R. Samy [3]).

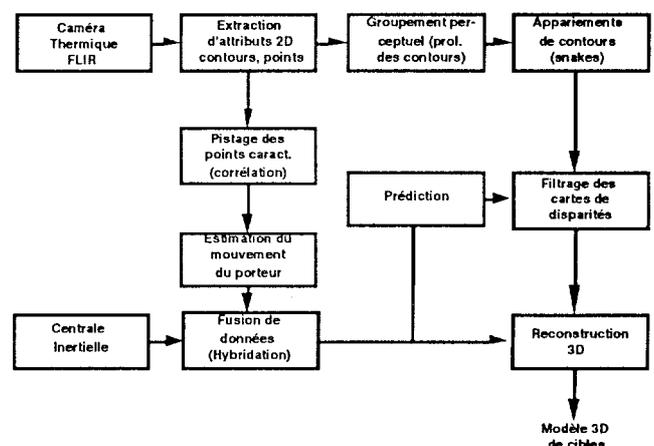


Fig 1 Diagramme général des traitements.



## II - MODELE GEOMETRIQUE DE LA STEREOVISION LATERALE

En stéréovision passive on distingue généralement deux approches: la stéréovision latérale et la stéréovision axiale [4]

La probabilité d'occurrence d'appariements erronés est plus faible en stéréovision axiale qu'en stéréovision latérale. Nous nous intéressons néanmoins à cette dernière pour son intérêt opérationnel dans le contexte d'images IR aériennes; le système de navigation de l'aéronef porteur permettant de résoudre un grand nombre d'incertitudes.

La stéréovision sur des séquences d'images IR en mouvement latéral, nous amène à formuler les équations régissant les mouvements observés dans les images vis-à-vis du mouvement du porteur [5].

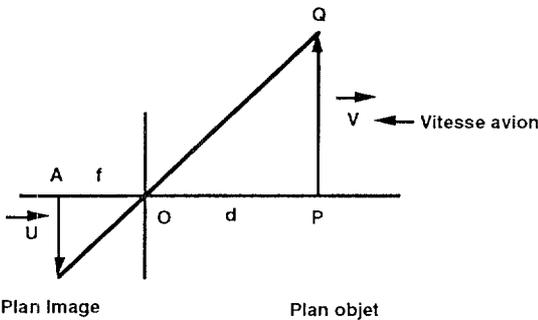


Fig2 Schéma d'acquisition de séquence d'images FLIR en mouvement latéral

En effet, si l'on considère un point mobile  $P_o$  de la scène et le point  $P_i$  de l'image correspondant, on peut leur affecter respectivement des vitesses  $V_o$ , dans la scène, et  $V_i$  dans le plan image.

A un ensemble de points mobiles de la scène, on peut associer un ensemble de points dans l'image caractérisé par un champ de vitesse.

Si la scène survolée par l'avion ne contient pas d'objets mobiles, on peut alors interpréter le champ de vitesse observé dans le plan image.

En utilisant les relations géométriques élémentaires, on obtient :

$$\frac{\mathbf{d}}{\mathbf{f}} = -\frac{\mathbf{V}}{\mathbf{u}} \quad (1)$$

d'où 
$$\mathbf{u} = -\mathbf{f} \frac{\mathbf{V}}{\mathbf{d}} \quad (2)$$

Si la caméra se déplace avec une vitesse constante  $V$ , on obtient :

$$\mathbf{u} = -\frac{\mathbf{k}}{\mathbf{d}} \quad (3)$$

Cette équation combinée à l'équation du flot optique simplifiée à un seul axe fournit :

$$\mathbf{u} = -\frac{\mathbf{I}_t}{\mathbf{I}_x} \text{ avec : } (\mathbf{I}_x = \frac{\partial \mathbf{I}}{\partial X}, \mathbf{I}_t = \frac{\partial \mathbf{I}}{\partial t}) \quad (4)$$

$$\text{d'où } \mathbf{d} = \frac{\mathbf{K I}_x}{\mathbf{I}_t} \quad (5)$$

Connaissant  $K$ , on peut alors reconstruire la géométrie de la scène.

## III - DETECTION ROBUSTE DE CONTOURS

L'optimalité des opérateurs de contours est une notion relative et certainement non vérifiée dans des images complexes telles que les images aériennes IR où l'on retrouve l'imbrication: détection, estimation.

Il convient alors d'introduire des techniques de détection-estimation couplées pour détecter des régions non homogènes dans un voisinage spatial défini et ré-estimer les moyennes locales afin d'augmenter le contraste d'une frontière pour aboutir à une plus grande probabilité de détection.

La théorie des estimateurs robustes devient très courante dans le traitement des images en raison de leur tolérance aux occlusions et perturbations de données erronées. Si l'on reformule l'estimation de paramètres associés à un modèle à partir de données bruitées, on peut définir le critère à minimiser comme une fonction non-linéaire  $\rho$  (au lieu de la norme L2) des écarts entre les mesures et une fonction des paramètres.

$$\Psi(r) = d\rho(r) / dr \quad (6)$$

La minimisation du critère conduisant à la résolution d'un système d'équations non-linéaires (au lieu d'un système linéaire quand il s'agit d'un critère quadratique L2) on est donc amené à introduire un algorithme itératif du type IRLS.

$$S_j = \frac{\Psi(r_j)}{r_j} \quad (7)$$

Ce dernier conduit à une estimée  $X$  qui s'exprime en fonction d'une matrice  $B$ , des données  $X$  et d'une matrice diagonale  $W$  représentant les résidus.

$$\hat{X} = (B^T W B)^{-1} B^T W d \quad (8)$$

En effet, à partir d'un calcul simple de gradient défini comme la différence de moyennes locales pondérées, on initialise un calcul itératif basé sur l'algorithme IRLS pour ré-estimer ces moyennes locales dont les différences définissent une composante du gradient spatial. L'affinement de l'estimation apporté par l'algorithme IRLS nous a permis d'obtenir une image d'amplitude de gradient ayant une dynamique plus grande, donc un rehaussement de contours, que celle produite par l'opérateur de Canny-Deriche.

Ce nouvel opérateur de détection de contours agit comme un opérateur multi-résolution, la différence provient du fait que cette analyse multi-résolution n'est pas régulière et définie a priori par des filtres passe-bande (Ondelettes, QMF) optimisés dans le domaine spectrale (analyse de Fourier) mais plutôt une organisation spatiale basée sur une analyse statistique des points (pixels) en fonction de leur niveau de gris.

#### IV - ESTIMATION DE MOUVEMENT POUR LA DETECTION D'OBJETS PRESENTANT UN RELIEF LOCAL

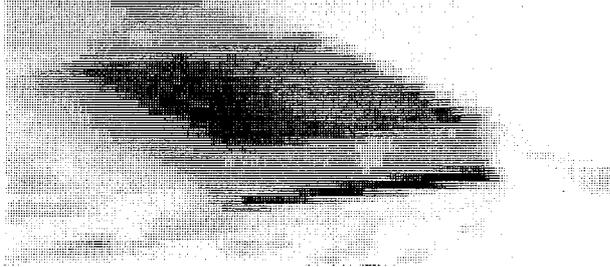
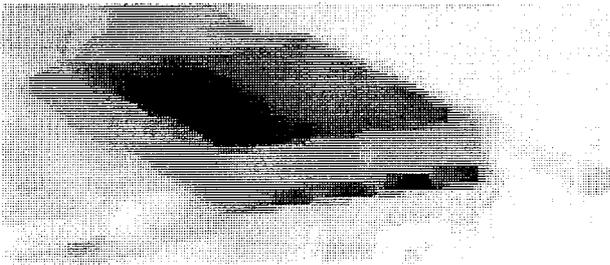


Image d'origine



Filtrage pondéré classique.



Filtrage robuste de Lorentz

Fig 3 Filtrage robuste



Extraction de contour Canny-Deriche



Extraction robuste Lorentz

Fig4 Extraction robuste de contours.

Afin d'estimer un champs de déplacements fiable et précis dans une séquence d'images IR aéroportée, nous utilisons l'approche classique de la mise en correspondance de points caractéristiques.

Après étude des différents extracteurs de points caractéristiques existants (Haralick, Beaudet, ... ), l'opérateur de Kanadé s'est avéré le plus robuste dans notre contexte tant au niveau de la pertinence des points sélectionnés que la stabilité temporelle de ces derniers.

L'opérateur de Tomasi-Kanadé [1] est basé sur le calcul des valeurs propres de la matrice G

$$G = \int_w gg' \omega \cdot dA \quad (9)$$

$\omega$  fonction gaussienne  
 $g$  gradient local  
 $W$  voisinage du point

Les estimations locales du gradient s'opèrent par application des masques dérivateurs de Haralick [2].

Les valeurs propres de la matrice G en un point caractéristique doivent vérifier la condition suivante:

$$\min(\lambda_1, \lambda_2) > \lambda_s \quad (10)$$

Les points caractéristiques sont extraits dans chaque image par cet opérateur puis sont mis en correspondance via une fonction de corrélation robuste. Une interpolation locale de la fonction de corrélation permet d'obtenir une estimation sub-pixélique du déplacement en chaque point sélectionné.

Le champs de déplacements estimé par cette méthode va être utilisé pour estimer les paramètres global du mouvement 2D engendré dans le plan image par le déplacement du porteur.

Dans l'hypothèse de planarité de la scène observée, le mouvement  $(dx, dy)$  induit dans l'image par le déplacement du porteur s'écrit sous la forme suivante polynomiale du second degré à l'aide de huit paramètres:

$$dx = a_0 + a_1X + a_2Y + a_6XY + a_7X^2$$

$$dy = a_3 + a_4X + a_5Y + a_6Y^2 + a_7XY$$

L'estimation de ces huit paramètres s'effectue à partir des observations locales des déplacements en chaque point caractéristiques par minimisation globale à l'aide d'une méthode de moindres carrés itératifs pondérés (IRLS) afin de rejeter les points présentant un relief (points minoritaires par rapport aux amers au sol).

La grande précision du recalage pixélique des images permet de recalibrer les données inertielles du porteur via un filtre d'hybridation des données.

Les paramètres du mouvement 2D induit dans l'image par le déplacement du porteur dépendent



directement de l'équation supposé du plan 3D observé par la camera .

L'approche consiste alors à estimer localement sur une région entourant un bâtiment les paramètres du mouvement observé et l'équation du plan du sol correspondant . Un calcul similaire est ensuite effectué sur les points du toit de l'immeuble; ce toit définissant un plan 3D différent de celui du sol. Une fois les équations de ces deux plans, sol et toit, estimées, on calcule l'élévation locale du toit par rapport au sol par simple différence d'altitude en un point situé au centre du toit.

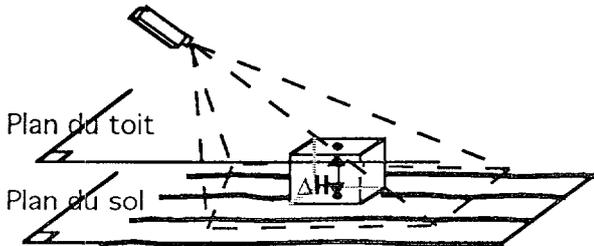


Fig5 Estimation locale de relief.

Il est à noter que les deux plans ne sont pas nécessairement parallèles. Le relief local ainsi que l'inclinaison de la face du toit observée sont donc pris en compte.

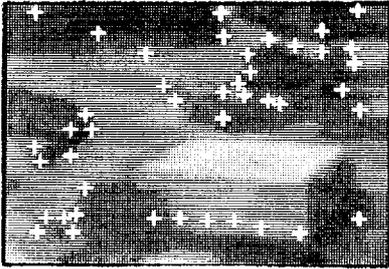


Fig6 Points caractéristiques dans le plan du sol

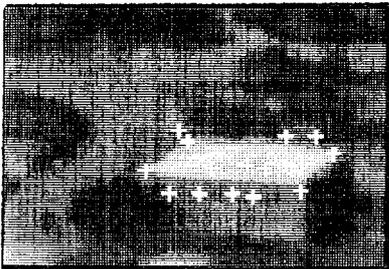


Fig7 Points caractéristiques dans le plan du toit.

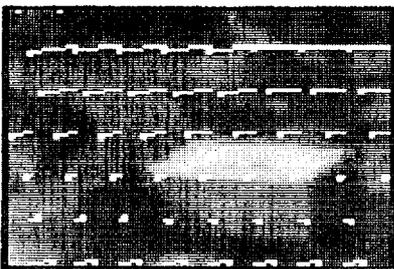


Fig8 Flot optique modélisé au sol.

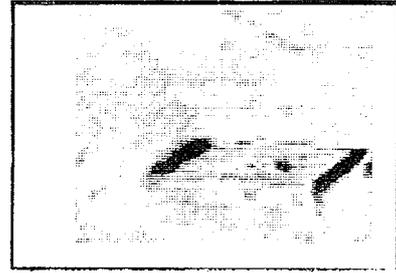


Fig9 Mise en évidence du relief par différence d'images après compensation du mouvement du porteur au niveau du sol.

Les deux équations des plans, sol et toit, fournissent une estimation de l'élévation au centre de l'immeuble de l'ordre de quelques mètres.

## V CONCLUSION ET PERSPECTIVES

Dans le cadre de la reconnaissance aérienne optronique IR, nous avons présenté une chaîne de traitements de stéréovision latérale mettant en oeuvre des estimations robustes pour réaliser la reconstruction 3D.

Cette reconstruction n'a pas été possible à très longue distance, seule une détection du relief est fiable.

## VI REMERCIEMENTS

Les travaux présentés dans cet article ont été réalisés dans le cadre de l'étude DRET (contrat n° 93/327) en collaboration avec l'équipe de l'ONERA/DES. Nous les remercions pour les images IR qui nous ont permis de réaliser ces expérimentations.

## VII BIBLIOGRAPHIE

[1] Tomasi-Kanadé "Shape and Motion from Image Streams: a factorisation method Part3: Detection and tracking of point features" April 1991 CMU-CS-91-132 Carnegie Mellon University Pittsburg

[2] Zuniga-Haralick "Corner detection using the facet model" Proc. Conf. on Pattern Recognition Image Analysis 1983 30-37

[3] JF Bonnet R. Samy "Développement de contours actifs robustes et adaptatifs. Application à la poursuite d'objets dans une séquence d'images IR. GRETSI 1995 Juan Les Pins

[4] V.Rodin, A.Ayache "Stéréovision axiale: modélisation et calibrage du système de prises de vues, reconstruction 3D d'objets naturels" Revue Traitement du Signal du GRETSI Vol 11N°5 1994

[5] R.Samy, D.Duclos "Pyramidal MRF models for optical flow estimation applied to target detection" SPIE Aerosense 94, Orlando USA April 1994