

Interprétation et Restitution automatique de bâtiments en milieu péri-urbain

Tuan DANG

Laboratoire MATIS, Institut Géographique National /DT/SR, 2 av.Pasteur, 94160 St.Mandé
Email: dang@cogit.ign.fr

Olivier JAMET

Laboratoire MATIS, Institut Géographique National /DT/SR, 2 av.Pasteur, 94160 St.Mandé

Henri MAITRE

Département Images, E.N.S.T, 46 rue Barrault, 75014 Paris

Résumé

Nous décrivons dans cet article une approche efficace pour améliorer la restitution automatique des bâtiments à partir d'un couple stéréo de photographies aériennes. Cette approche fait appel à l'utilisation des modèles de surface sur les disparités obtenues par corrélation automatique classique. Pour cela, des techniques d'analyse d'images ont été employées conjointement avec la corrélation automatique pour délimiter les domaines de définitions des surfaces analytiques.

Mots clés: Corrélation, Segmentation, Modèles de surface,
Analyse d'image, Groupement Perceptuel.

Abstract

In this paper, we describe an efficient approach to improve stereo reconstruction of buildings from a stereo pair of aerial images. We use analytic models of surface depth to remove noises in the disparity map obtained by conventional stereo algorithm. In order to achieve this, image analysis techniques have been used to find boundaries of analyticity of models.

Key words: Stereo algorithm, Segmentation, Models of surface,
Image Analysis, Perceptual grouping.

1) But et Contexte

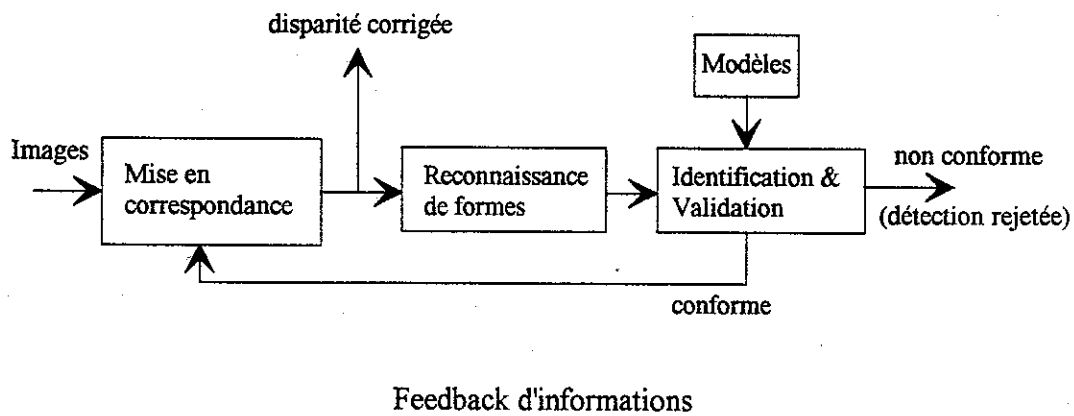
Cette recherche est entreprise par l'I.G.N au sein du laboratoire MATIS pour essayer de trouver des méthodes de restitution automatique ou semi-automatique de l'information topographique afin de réduire les délais de production de ces informations. En effet, à l'heure actuelle la saisie de ces informations est faite manuellement sur des clichés analogiques à l'échelle 1:30000 à l'aide de restituteurs analytiques; ce qui nécessite un temps considérable.

Dans ce papier, nous décrivons les travaux que nous avons entrepris depuis un an sur le problème de la restitution automatique des bâtiments à partir d'un couple stéréo de photographies aériennes à moyenne échelle, numérisées avec un pas d'échantillonnage au sol de l'ordre du mètre. La restitution des bâtiments est confrontée particulièrement à plusieurs difficultés, d'une part la complexité et la variabilité des structures empêchent l'emploi seul des techniques classiques de reconnaissance de formes; d'autre part les ruptures de pentes, les parties cachées et le manque de contraste constituent de véritables obstacles à l'utilisation de la corrélation automatique classique. Nous exposons, dans ce qui suit, notre démarche et les différentes techniques utilisées ainsi que les résultats obtenus, puis nous concluons sur les travaux futures qui restent à réaliser.

2) Démarche proposée

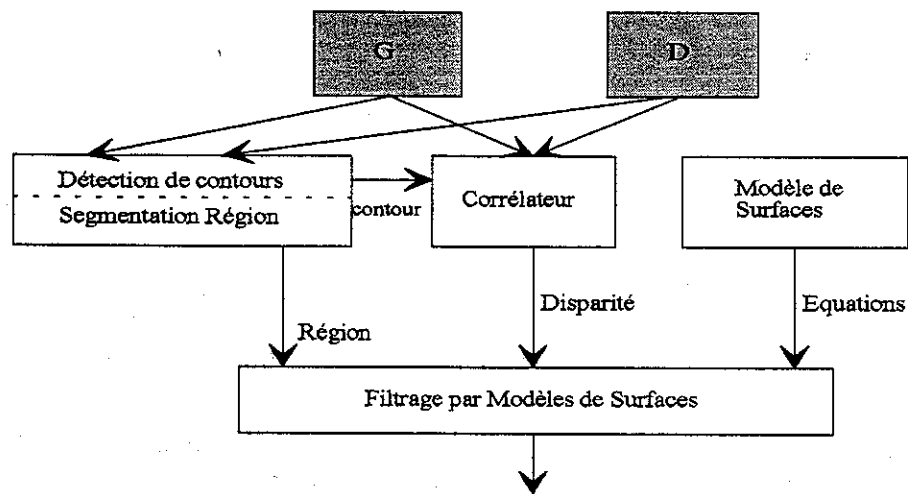
La solution que nous proposons ici consiste à compléter les travaux utilisant des modèles de surfaces analytiques dans le processus de filtrage des disparités [Eastman 87, Hoff 89, Maître 92] par l'utilisation d'une reconnaissance de formes. Eastman et Waxman ont utilisé des modèles plans et quadratiques sur des petits voisinages dans un repère de l'image pour guider la corrélation par une estimation locale de la disparité et opérer simultanément un filtrage des parallaxes grâce aux modèles. Ces auteurs travaillent directement sur l'image des disparités. Hoff et Ahuja ont utilisé des modèles plans et quadratiques dans un algorithme de type pyramidal pour guider la corrélation. A la différence de Eastman et Waxman, ils travaillent dans un système de coordonnées de l'objet. L'approche proposée par Maître et Luo consiste à faire coopérer les informations photométriques avec les modèles plans et quadratiques des surfaces afin de filtrer et d'interpoler les disparités obtenues par des algorithmes de mise en correspondance proposés par Wu [Wu 88] et par Grimson [Grimson 85]. L'information photométrique est traduite sous forme de régions, obtenues par un algorithme de segmentation. Le problème de cette approche est que les discontinuités radiométriques ne correspondent pas toujours à celles du relief. Ceci peut produire certains artéfacts, auxquels s'ajoutent les imperfections de la segmentation. L'approche que nous proposons ici consiste à compléter cette méthode par l'utilisation de reconnaissance de forme, qui nous permettra de

corriger les artefacts causés par la segmentation. Notre approche peut être résumée par le schéma suivant:

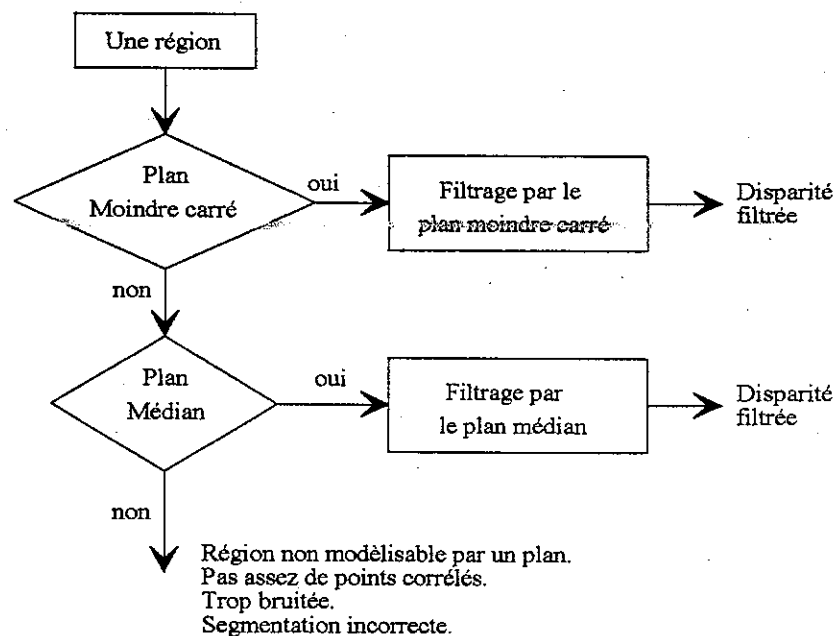


Mise en correspondance:

Le rôle de cette partie est de nous fournir une carte de disparités filtrées par des modèles de surfaces [Maître 92]. Pour cela, une première estimation des parallaxes est obtenue par corrélation automatique. L'algorithme que nous développons consiste à corrélérer à la fois les contours et les imageries de niveau de gris. L'utilisation en plus des contours nous permet d'avoir une meilleure confiance dans les points homologues trouvés. En effet nous imposons que l'homologue d'un point contour soit aussi un point contour. Pour limiter le domaine de recherche, les images sont rééchantillonnées en épipolaires. Les contours sont obtenus grâce au filtrage optimal récursif de Canny-Deriche [Deriche 87]. Par ailleurs, nous avons utilisé une technique de correction et de propagation des disparités le long des contours similaire à celle proposée par Mohan et al [Mohan 89]. Cette technique repose sur le fait que les disparités doivent varier linéairement le long des segments de droite détectés. Lorsqu'une première estimation des parallaxes est obtenue, les modèles plans de surface sont ensuite testés à l'intérieur des régions de l'image de gauche qui sert de référence lors du calcul des disparités. Ces régions sont obtenues par une méthode de segmentation proposée par Suk et al [Suk 83]. Nous nous sommes limités au modèle plan de surface car il permet de décrire la plupart des structures faites de main d'homme. Le schéma suivant résume l'interaction des différentes informations durant cette étape:



Le filtrage par modèles de surfaces est illustré par l'organigramme suivant:



Le filtrage par modèle plan a été amélioré par rapport à celui proposé par Luo [Luo 91] en faisant intervenir des plans médians, qui nous permettent de filtrer des bruits non-gaussiens. D'autre part, le filtrage par plan médian est plus robuste que celui par plan aux moindres carrés dans le cas des surfaces horizontales qui représentent, pour les scènes en milieu péri-urbain, une part non négligeable des toits de bâtiments.

Pour un **plan au moindre carré**, le test de validation est le suivant:

- 1) $\text{Card}(\text{Région}) > N$.
- 2) Le nombre de points corrélés est supérieur à 3.
- 3) Ecart quadratique moyen par rapport au plan candidat est inférieur à ϵ^2 .
- 4) Gradient de disparité est inférieur à 2.

Pour un **plan médian**, le test est:

- 1) $\text{Card}(\text{Région}) > N$.
- 2) Le nombre de points corrélés est supérieur à 3.
- 3) Ecart moyen par rapport au plan médian est inférieur à ϵ .

N et ϵ sont des seuils fixés à l'entrée du programme.

Reconnaissance de formes:

Le rôle de cette étape est d'apporter au processus de mise en correspondance, des informations complémentaires sur la forme des bâtiments à restituer. Ces informations vont d'une part, "aider" l'algorithme stéréo à mieux corréler les bâtiments détectés, d'autre part elles vont servir à corriger les artéfacts de la segmentation par un ajustement des frontières des régions. L'identification et la validation des bâtiments détectés font appel à une comparaison avec des modèles génériques de bâtiments. Les formes reconnues comme étant conformes aux modèles sont réinjectées au processus de mise en correspondance. De cette façon, l'algorithme de mise en correspondance dispose d'informations complémentaires sur les bâtiments détectés. Ces bâtiments seront ensuite reappariés pour donner une nouvelle carte de disparité.

Le modèle utilisé décrit les bâtiments comme des assemblages de primitives (segments de droite, rectangles...) disposées orthogonalement les unes par rapport aux autres. Ceci nous conduit à détecter ces primitives géométriques nécessaires à la reconnaissance des bâtiments. Pour cela nous faisons appel à des techniques de détection de formes par groupement perceptuel. Le groupement perceptuel englobe six types de groupement [Lowe 85]: groupement de proximité, groupement de similarité, groupement de continuation, groupement de fermeture, groupement de symétrie et groupement de familiarité. En fait, ces groupements reflètent un mécanisme de reconnaissance de formes chez l'être humain, qui est capable, à la vue des objets présents dans une image, d'établir des groupements d'objets présentant des propriétés particulières grâce à ce mécanisme de perception. Ces mécanismes ont été étudiés et proposés par l'école des Gestaltistes dès les années 1920-1930 [Wertheimer 58]. En fait, grâce au groupement perceptuel, nous pouvons extraire les primitives pertinentes pour l'interprétation et réduire en même temps la quantité des données et tenir compte du bruit. En général, le problème du groupement perceptuel est lié à trois sous-problèmes. Premièrement, nous devons résoudre le problème de choix et de représentation des primitives à partir desquelles nous allons extraire les objets recherchés. Deuxièmement, nous devons réfléchir sur le type de groupement à faire sur ces primitives. Enfin nous devons résoudre le problème de choix d'un groupement parmi les groupements possibles.

Nous avons choisi les segments de contour comme primitives de travail, qui sont obtenues par chaînage et approximation polygonale des pixels de contour binaire. L'approximation polygonale est réalisée grâce à l'algorithme de Wall et Danielsson [Wall 84].

Ces contours binaires proviennent d'une part de l'application du détecteur de Canny-Deriche sur l'image initiale, et d'autre part de la détection sur l'image de parallaxe. L'utilisation des contours détectés sur l'image de disparité nous permet de tirer profit de l'information du relief fournie par l'étape de mise en correspondance. En effet la détection des ruptures de pente permet de localiser des structures susceptibles de représenter les bâtiments. Quant à l'utilisation des contours détectés sur l'image initiale, elle permet d'apporter des informations complémentaires sur la forme des bâtiments détectés au processus de mise en correspondance.

Détection robuste de structures linéaires:

La détection des structures linéaires par groupement perceptuel a été étudiée récemment par un certain nombre d'auteurs [Weiss 86, Reynolds 87, Mohan 89, Lu 92]. Weiss et Boldt utilisent des critères géométriques et photométriques pour grouper des segments constituant les noeuds d'un graphe, dont les arcs sont définis par une relation binaire entre noeuds. La fusion est effectuée en examinant tous les chemins menant d'un noeud initial à un noeud final appartenant à un domaine de recherche donné. Les noeuds du chemin vérifiant des critères géométriques et photométriques sont ensuite fusionnés. Le problème de cette approche est que le coût de fusion peut être très élevé, lorsqu'il y a un nombre important de segments. Reynolds et Beveridge ont utilisé une approche similaire, mais avec des critères uniquement géométriques. Cependant, ces auteurs ne font pas de fusion de segments. L'approche proposée par Mohan et Nevatia d'une part, par Lu et Aggarwal d'autre part, consiste à grouper itérativement les segments de même orientation qu'un segment S donné et appartenant à un voisinage de S . Les segments groupés sont ensuite remplacés par un seul segment, dont l'orientation est la moyenne des orientations pondérées de la longueur de chaque segment.

La solution que nous proposons dans cette partie, consiste à rendre l'approche de Lu et Aggarwal plus robuste, et à apporter un formalisme clair au groupement proposé. En effet, les approches de Mohan et Nevatia d'une part, de Lu et Aggarwal d'autre part, ne permettent pas de prolonger les segments. Par ailleurs, le groupement proposé par ces auteurs manque de robustesse, en cas de présence de nombreux segments "bruits", qui sont souvent très courts (de 2 à 4 pixels), dont l'orientation peut faire diverger fortement l'orientation finale de l'orientation du segment le plus long (parmi les segments groupés). Le groupement proposé ici permet de prolonger les segments et de régler la robustesse du processus.

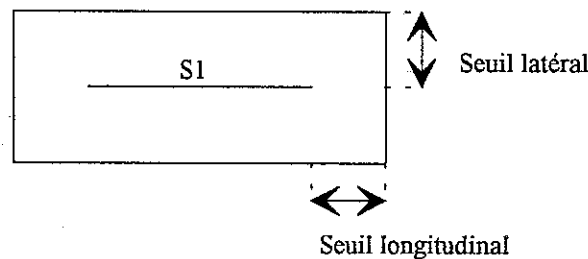
Dans ce qui suit, nous exposons principalement la détection des structures linéaires qui vont servir de base à la détection des rectangles susceptibles de représenter les bâtiments. Par ailleurs, la détection des structures linéaires nous permet de réduire la quantité des données à traiter au niveau supérieur. La détection des structures linéaires est effectuée en deux étapes. La première étape consiste à linéariser les segments jointifs afin de combler certaines lacunes de l'algorithme d'approximation polygonale. La deuxième étape est le groupement des

segments proches et colinéaires via une relation de colinéarité \mathcal{R}_c définie de la manière suivante:

Soit S_1 et S_2 deux segments de droite avec $S_2 = [AB]$

$$S_1 \mathcal{R}_c S_2 \Leftrightarrow \begin{cases} 1) \text{ Rec}(S_1) \cap S_2 \neq \emptyset \\ 2) |\cos(S_1, S_2)| \geq \varepsilon \\ 3) d(A, S_1) \leq s_{la} \text{ et } d(B, S_1) \leq s_{la} \end{cases}$$

où: $\text{Rec}(S_1)$ désigne un voisinage rectangulaire englobant le segment S_1 ,
 ε étant un seuil fixe, s_{la} désigne un seuil latéral.
 $d(A, S_1)$ et $d(B, S_1)$ étant les distances de A à S_1 et de B à S_1 .



$\text{Rec}(S_1)$ est défini par deux seuils: seuil latéral et seuil longitudinal qui sont fixés pour l'ensemble des segments de l'image.

Remarque: \mathcal{R}_c n'est ni transitive, ni symétrique.

Le groupement est réalisé de la façon suivante. Pour un segment s donné, nous cherchons l'ensemble G_s défini par:

$G_s = \{s' \in I / s \mathcal{R}_c s'\}$, les s' sont orientés dans le **même sens** que s
 où I désigne l'ensemble des segments de l'image.

Afin de réduire le domaine de recherche pour un segment s donné, nous ne considérons que les segments candidats qui ont vérifié la condition 1 de \mathcal{R}_c . En effet, l'ordre des conditions est classé suivant le nombre d'opérations à effectuer. Ceci nous permet d'accélérer le traitement par une réduction du nombre de calculs à faire pour vérifier les conditions de \mathcal{R}_c .

Ensuite G_s est remplacé par S_f défini par:

$$1) (\bar{i}, \bar{S}_f) \text{ est défini par: } \cos(\bar{i}, \bar{S}_f) = \frac{\sum_{s \in G_s} \|\bar{s}\|^n \cdot \cos(\bar{i}, \bar{s}')}{\sum_{s \in G_s} \|\bar{s}\|^n} \text{ et } \sin(\bar{i}, \bar{S}_f) = \frac{\sum_{s \in G_s} \|\bar{s}\|^n \cdot \sin(\bar{i}, \bar{s}')}{\sum_{s \in G_s} \|\bar{s}\|^n},$$

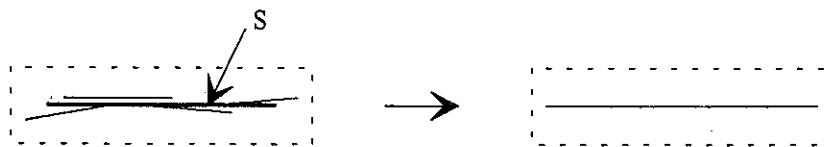
\bar{i} désigne l'axe des abscisses, $(\bar{i}, \bar{s}') \in [-\pi/2, \pi/2]$, $n \in \mathbb{N}^*$

2) \bar{S}_f passe par le barycentre des segments de G_s

3) Les deux extrémités de S_f sont les deux projetés les plus éloignés parmi les projetés sur S_f

Ici, il nous semble plus judicieux de faire passer S_f par le barycentre des segments affectés d'un poids égal à leur longueur, que de faire passer S_f par l'isobarycentre des segments comme faisaient Lu et Aggarwal. En fait, il est légitime de considérer que plus un segment est long, plus son poids est important. Le fait de pondérer l'orientation des segments participant à G_s par une fonction de la longueur des segments ayant une croissance polynomiale, nous permet de diminuer l'influence des segments "bruits". Ici, nous avons choisi un monôme de degré n . Le choix de n permet de jouer sur la robustesse du groupement. Par ailleurs, l'intervention d'un seuil longitudinal nous permet de prolonger le segment final lors de chaque groupement.

La figure suivante illustre le type de groupement que nous souhaitons réaliser:



groupement de proximité et de similarité

En ce qui concerne la linéarisation de segments jointifs, il s'agit de réaliser le groupement suivant:



Linéarisation de segments jointifs

3) Résultats

Nous montrons dans ce qui suit, quelques résultats illustrant les différentes étapes de notre démarche:

Les résultats ci-dessous représentent les données en entrée et en sortie de l'étape de mise en correspondance:

En entrée, nous fournissons au processus un couple stéréo en géométrie épipolaire:

Couple stéréo numérisé au pas de 1.2m par pixel:



Image de gauche



Image de droite

En sortie, nous obtenons une image de disparité corrigée par des modèles plans:

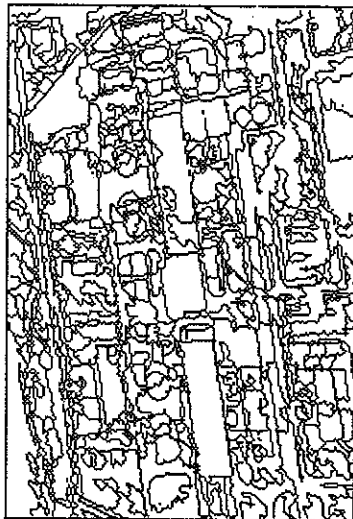
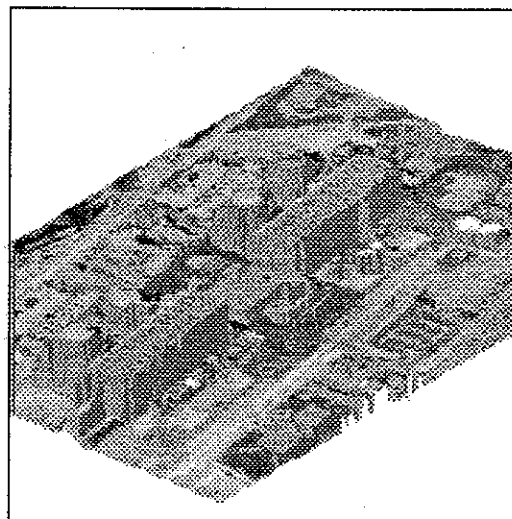


Image de gauche

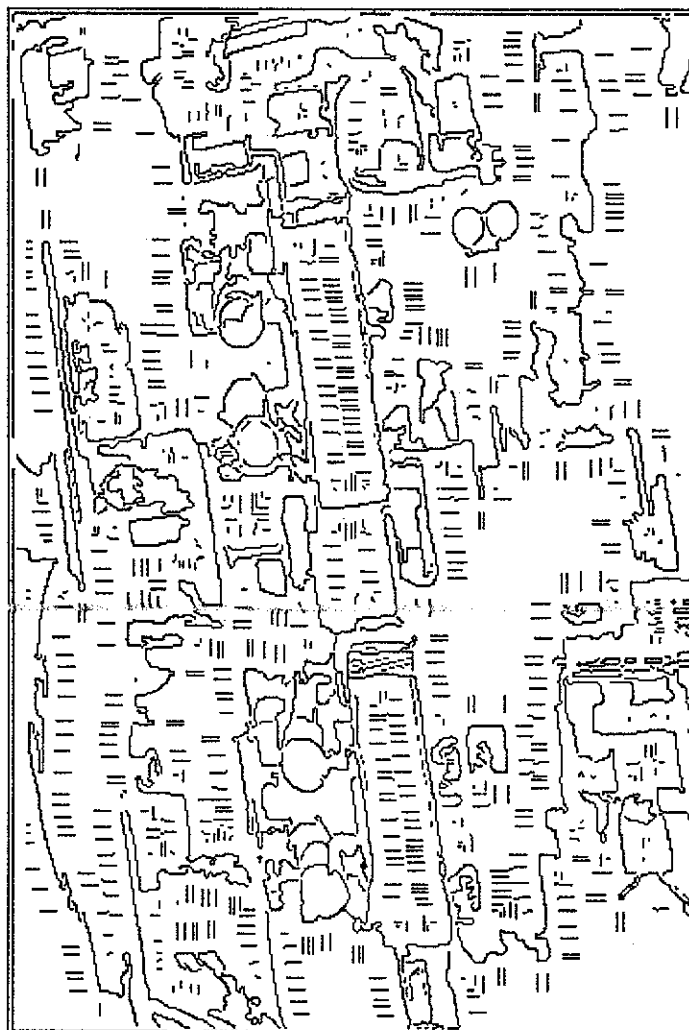
(segmentée par Suk: T1=10, T2=10, T3=100)



Carte de disparité filtrée

Sur l'image de disparité, on note la présence de quelques artéfacts liés à l'imperfection de la segmentation et au fait que les discontinuités radiométriques ne correspondent pas à celles du relief.

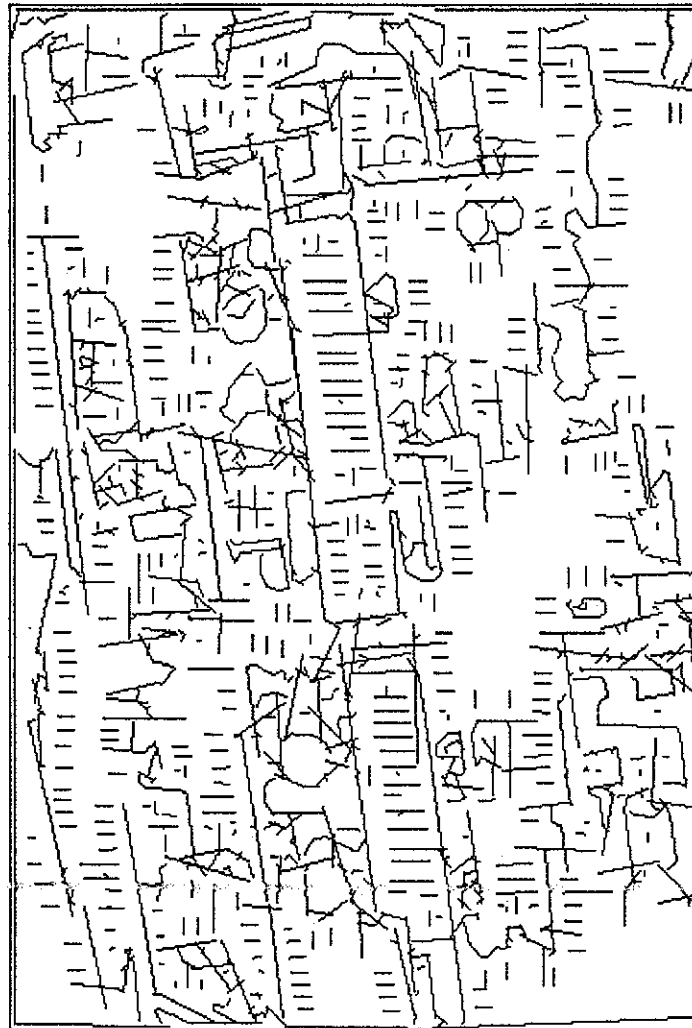
Les résultats suivants illustrent la détection des structures linéaires qui vont servir à la détection des primitives géométriques nécessaires à la reconnaissance des bâtiments:



Contour détecté sur l'image de disparité

L'utilisation de ces contours permet d'une part, de détecter les ruptures de pente qui sont utiles pour repérer des structures susceptibles de représenter les bâtiments, d'autre part, elle nous permet de ne corriger des imperfections de la segmentation qu'aux zones d'intérêt de l'image. De cette manière, le traitement sera plus rapide grâce à une réduction des données.

Nous montrons dans ce qui suit, les structures linéaires détectées sur l'image ci-dessus:



seuil latéral = 4pixels, seuil longitudinal = 8pixels, tolérance=20°, n=1

Ce résultat montre encore une fois, des possibilités du groupement perceptuel dans la détection des structures artificielles, en occurrence, des structures linéaires. Ces structures vont servir à la détection des primitives de plus haut niveau, susceptibles de représenter les bâtiments.

4) Conclusion

Ces résultats intermédiaires montrent encore une fois l'efficacité de l'approche filtrage des disparités par modèles de surfaces analytiques. Toutefois, ce type de méthode introduit quelques artéfacts du fait qu'il repose sur l'utilisation d'une partition de l'image en régions radiométriquement homogènes. Or les discontinuités radiométriques ne correspondent pas toujours aux discontinuités du relief. Ainsi, nous pensons qu'il est indispensable d'utiliser les outils de l'analyse d'image qui vont "aider" la stéréo à "voir" certaines parties cachées en lui fournissant des informations complémentaires sur la forme des objets à restituer. Cette coopération entre une technique de la stéréovision classique et celle de la reconnaissance de forme devrait nous donner un meilleur résultat du fait de la complémentarité de chaque

technique. Le but que nous poursuivons dans les prochains travaux, est de détecter les rectangles susceptibles de représenter les bâtiments, afin de corriger les artéfacts de la segmentation, mais aussi de corriger la carte de disparité par un apport d'informations supplémentaires au processus de mise en correspondance.

Références:

- [Deriche 87] R. Deriche, "Using Canny's criteria to derive a recursively implemented optimal edge detector", International Journal of Computer Vision, May, 1987.
- [Eastman 87] Roger D. Eastman and Allen M. Waxman, "Using Disparity Functionals for Stereo Correspondance and Surface Reconstruction", C.V.G.I.P, Vol.39, pp.73-101, 1987
- [Grimson 85] Grimson W.E.L, "Computational Experiments with a Feature Based Stereo Algorithm", I.E.E.E. P.A.M.I, Vol.7, pp.17-34, 1985.
- [Hoff 89] William Hoff, and Narendra Ahuja, "Surface from Stereo: Integrating Feature Matching, Disparity Estimation, and Contour Detection", I.E.E.E. P.A.M.I, Vol.11, N° 2, pp.121-136, Feb. 1989.
- [Lowe 85] David G.Lowe, "Perceptual Organization and Visual Recognition", Kluwer Academic Publisher, Hingham MA 02043, USA, 1985.
- [Lu 92] H.Q.Lu and J.K.Aggarwal, "Applying Perceptual Organization to the Detection of man-made objects in non-urban scenes", Pattern Recognition, Vol.25, No.8, pp.835-853, 1992.
- [Luo 91] Wei LUO, "Utilisation des Modèles de Surfaces en Stéréovision", Thèse de l'Ecole Nationale Supérieure des Télécommunications de Paris, 1991.
- [Maître 92] Henri Maître and Wei Luo, "Using Models to Improve Stereo Reconstruction", I.E.E.E-P.A.M.I, Vol.14, No. 2, pp.269-277, February 1992.
- [Mohan 89] R.Mohan, G.Medioni and R.Nevatia, "Stereo Error Detection, Correction, and Evaluation", I.E.E.E-P.A.M.I, Vol.11, No. 2, pp.113-120, Feb. 1989.
- [Mohan 89] R.Mohan, and R.Nevatia, "Using Perceptual Organization to Extract 3-D Structures", I.E.E.E-P.A.M.I, Vol.11, No. 11, pp.1121-1139, Nov. 1989.
- [Reynolds 87] G. Reynolds and J.Ross Beveridge, "Searching for Geometric Structure in Images of Natural Scenes", Proceeding: Image Understanding Workshop, Vol. 1, Feb. 1987.
- [Suk 83] M. Suk and S.M. Chung, "A New Image Segmentation Technique Based on Partition Mode Test", Pattern Recognition, Vol. 16, No. 5, pp. 469-480, 1983.
- [Wall 84] Wall K., Danielsson P.E., "A fast sequential method for polygonal approximation of digitized curves", C.V.G.I.P, 28, 1984, pp.220-227.
- [Weiss 86] R.Weiss and M.Boldt, "Geometric Grouping Applied to Straight Lines", Proceedings: CVPR, pp. 489-495, June 22-26, 1986, Miami Beach, Florida.
- [Wertheimer 58] M.Wertheimer, "Principles of Perceptual Organization", Readings in Perception, D.Bearslee and M.Wertheimer Eds, Princeton, 1958.
- [Wu 88] Wu Y.F and Maître Henri, "A New Dynamic Programming Method for Stereo Ignoring Epipolar Geometry", 9th ICPR, Rome 1988, pp.79-81.