

Modélisation et estimation robuste de la surface du sol en imagerie aérienne stéréoscopique

Thomas BELLI

Équipe Traitement des Images et du Signal (CNRS UPRES-A 8051)
Université de Cergy-Pontoise / ENSEA
6, av. du Ponceau 95014 Cergy-Pontoise Cedex, France
belli@ensea.fr

Résumé – Nous proposons une méthode pour estimer la surface du sol dans une scène urbaine à partir d’un Modèle Numérique d’Élévation. Notre approche s’appuie sur une modélisation paramétrique de la surface du sol décomposée sur une base de fonctions harmoniques 2D. Sous l’hypothèse que les éléments du sur-sol correspondent à des données aberrantes, les M-estimateurs fournissent un cadre théorique pour l’estimation robuste des paramètres du modèle de surface. Nous présentons des résultats de reconstruction du sol sur différentes scènes urbaines.

Abstract – We propose a novel approach dedicated to the extraction of above-ground objects in urban scenes by means of a Digital Elevation Model. This method is based on the modeling of the terrain with a parametric surface decomposed on a Fourier basis of 2D harmonic functions. Under the assumption that above-ground objects are outliers, M-estimators supply a theoretical framework to obtain a robust estimation of the ground surface parameters. We present results obtained on several urban scenes.

1 Introduction

De nombreux travaux portent sur la modélisation de scènes urbaines à partir de données 3D, la plupart d’entre eux étant consacrés à la modélisation du bâti. Un aspect important réside cependant dans la séparation des éléments de la scène en deux classes: sol et sur-sol. Cette étape de classification s’inscrit dans une stratégie hiérarchique de focalisation. Elle est généralement suivie par une classification des éléments du sur-sol en deux classes: bâti et végétation. Des modèles spécifiques peuvent alors être appliqués aux bâtiments comme le modèle de mélange multi-plans proposé par Cord et Declercq [1].

Deux types d’approches sont généralement suivies pour l’identification des éléments du sur-sol: *locale* et *globale*.

Le premier se base uniquement sur une information *locale* et regroupe les méthodes exploitant essentiellement les transitions altimétriques locales présentes dans le Modèle Numérique d’Élévation (MNE). Une ouverture morphologique [2] ou le filtre morphologique ”dual rank” proposé par Eckstein et Munkelt [3] permet d’obtenir une approximation du Modèle Numérique de Terrain (MNT) et donc l’extraction du sur-sol. Ces approches nécessitent de connaître la taille maximale d’un élément du sur-sol afin de définir la taille de l’élément structurant. Baltsavias et al. [4] proposent de détecter les éléments du sur-sol en recherchant les lignes de niveau fermées.

Le second type de méthodes s’appuie sur une modélisation *globale* à l’instar du modèle markovien proposé par Baillard [5]. Cette approche implique la minimisation d’une énergie globale à toutes les cliques, mais elle intègre toutefois la différence d’altitude entre deux cliques voisines dans la probabilité *a priori*.

Notre approche se base sur une modélisation globale de

la surface du sol par une fonction paramétrique bidimensionnelle. Un MNE obtenu par mise en correspondance d’un couple d’images aériennes fournit un ensemble de points 3D [6]. Puis, une estimation robuste des paramètres du modèle s’appuyant sur les propriétés des M-estimateurs permet de reconstruire la surface du sol en écartant les points du sur-sol. Cette stratégie aboutit conjointement à une estimation de la surface du terrain et une détection des points du sur-sol.

2 Modélisation de la surface

Afin de tenir compte des variations de pente présentes dans une scène, nous modélisons la surface du sol par une fonction $z(x, y)$ avec $(x, y) \in \mathbb{R}^2$ se décomposant sur base de fonctions orthogonales $\phi_k(x, y)$:

$$z(x, y) = \sum_k c_k \phi_k(x, y) \quad (1)$$

Nous avons choisi de représenter la surface dans une base de fonctions harmoniques 2D. Pour un MNT de taille $T_x \times T_y$, la surface décomposée en une série de Fourier à l’ordre N s’écrit:

$$\begin{aligned} z_{\theta}(x, y) = & a_{0,0} + \sum_{k,l=0; k+l \neq 0}^N a_{k,l} \cos(2\pi(k\nu_x x + l\nu_y y)) \\ & + \sum_{k,l=0; k+l \neq 0}^N b_{k,l} \sin(2\pi(k\nu_x x + l\nu_y y)) \end{aligned} \quad (2)$$

avec les coefficients de Fourier

$$\theta = (a_{0,0}, a_{0,1}, b_{0,1}, \dots, a_{N,N}, b_{N,N})^T$$

et les fréquences fondamentales $\nu_x = 1/T_x$ et $\nu_y = 1/T_y$.

L'ordre N du modèle détermine le degré d'approximation de la surface et son choix sera discuté dans les sections suivantes.

3 Estimation robuste

Le MNE obtenu par reconstruction stéréoscopique fournit un ensemble de points 3D. Notre stratégie consiste alors à rejeter les points du sur-sol par une estimation robuste des paramètres du modèle de sol introduit dans la section 2.

Nous disposons d'un ensemble Ω de m points de coordonnées $(x_i, y_i, z^d(x_i, y_i))_{i=1,m}$ pour l'estimation des $P = 2(N+1)^2 - 1$ paramètres du modèle. Ces points doivent satisfaire l'équation (2), ce qui conduit à résoudre le système matriciel suivant :

$$\mathbf{z}^d = M\boldsymbol{\theta} \quad (3)$$

avec

$$M = \begin{pmatrix} 1 & C_{0,1}(x_1, y_1) & S_{0,1}(x_1, y_1) & \cdots & S_{N,N}(x_1, y_1) \\ 1 & C_{0,1}(x_2, y_2) & S_{0,1}(x_2, y_2) & \cdots & S_{N,N}(x_2, y_2) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & C_{0,1}(x_m, y_m) & S_{0,1}(x_m, y_m) & \cdots & S_{N,N}(x_m, y_m) \end{pmatrix}$$

$$\mathbf{z}^d = (z^d(x_1, y_1), z^d(x_2, y_2), \dots, z^d(x_m, y_m))^T$$

où nous définissons $C_{k,i}(x, y) = \cos(2\pi(kv_x x + lv_y y))$ et $S_{k,i}(x, y) = \sin(2\pi(kv_x x + lv_y y))$ pour simplifier l'expression de la matrice M . Le nombre de points $m = \text{card}(\Omega)$ est très supérieur au nombre P de paramètres à estimer, le système (3) est donc sur-déterminé.

Une estimation de $\boldsymbol{\theta}$ par minimisation au sens des Moindres Carrés (MC) fournit un résultat erroné. En effet, de nombreux points du MNE appartiennent à des éléments du sur-sol et doivent être considérés comme des aberrations. Ainsi, nous réalisons une estimation robuste au moyen d'un M-estimateur [7]. Cette approche consiste à minimiser une fonction ρ_c de l'erreur afin de limiter l'influence des erreurs importantes sur l'estimation. Nous cherchons donc le vecteur de paramètres qui minimise l'expression suivante :

$$\boldsymbol{\theta}_{\rho_c} = \arg \min_{\boldsymbol{\theta}} \sum_{i \in \{1, \dots, m\}} \rho_c(\epsilon_i, \boldsymbol{\theta}) \quad (4)$$

où $\epsilon_i, \boldsymbol{\theta}$ représente l'erreur sur l'axe (0z) entre les données z_i^d et le modèle de surface $z_i, \boldsymbol{\theta} = (M\boldsymbol{\theta})_i$ au point (x_i, y_i) :

$$\epsilon_i, \boldsymbol{\theta} = z^d(x_i, y_i) - z_{\boldsymbol{\theta}}(x_i, y_i)$$

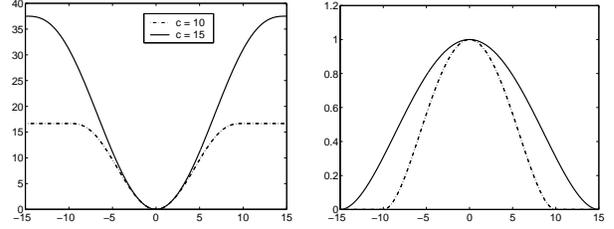
Nous employons le M-estimateur de Tukey donné par la fonction ρ_c suivante :

$$\rho_c(x) = \begin{cases} \frac{c^2}{6} \left(1 - \left[1 - \left(\frac{x}{c}\right)^2\right]^3\right) & \text{si } |x| \leq c \\ \frac{c^2}{6} & \text{sinon} \end{cases} \quad (5)$$

Le paramètre d'échelle c contrôle l'influence des points aberrants (du sur-sol) comme indiqué sur la figure 1.1. Par exemple, dans le cas du M-estimateur de Tukey, deux erreurs égales et supérieures à c ont la même contribution à l'erreur globale de l'équation (4).

À tout M-estimateur donné par une fonction ρ_c , symétrique, définie positive avec un unique minimum $\rho_c(0) = 0$, correspond une fonction de poids ω_c , où $\omega_c(x) = \frac{1}{x} \frac{\partial \rho_c(x)}{\partial x}$. Pour le M-estimateur de Tukey, ω_c s'écrit :

$$\omega_c(x) = \begin{cases} \left[1 - \left(\frac{x}{c}\right)^2\right]^2 & \text{si } |x| \leq c \\ 0 & \text{sinon} \end{cases} \quad (6)$$



1.1: Fonctions ρ_c .

1.2: Fonctions ω_c .

FIG. 1: M-estimateur de Tukey pour deux valeurs du paramètre d'échelle c .

Une propriété importante des M-estimateurs réside dans la simplification du processus d'optimisation. En effet, la minimisation de l'erreur globale dans l'équation (4) se traduit par une résolution itérée et repondérée au sens des MC jusqu'à convergence :

$$\boldsymbol{\theta}^{(k)} = \arg \min_{\boldsymbol{\theta}} \sum_{i \in \{1, \dots, m\}} \omega_c(\epsilon_i^{(k-1)}, \boldsymbol{\theta}^{(k-1)}) (\epsilon_i^{(k)}, \boldsymbol{\theta}^{(k)})^2 \quad (7)$$

La technique d'optimisation que nous employons consiste à initialiser les paramètres $\boldsymbol{\theta}$ du modèle par MC, puis à rejeter progressivement les points aberrants par une succession d'estimations robustes. À cette fin le paramètre d'échelle c du M-estimateur est initialisé à une valeur élevée et décroît jusqu'à une valeur limite c_l qui correspond à la hauteur minimale des points du sur-sol. Cette approche est similaire à une stratégie de type Graduated Non-Convexity. En effet, le M-estimateur de Tukey est quadratique pour un rapport x/c faible et devient moins convexe pour des valeurs croissantes de ce rapport.

Notons que la théorie des statistiques robustes suggère que, sous certaines conditions, l'estimation avec un M-estimateur est possible jusqu'à 50% de points aberrants. Notre approche est donc susceptible de fournir de bons résultats sur de nombreuses scènes urbaines, car le taux de recouvrement du sur-sol est généralement inférieur à cette limite.

Afin d'illustrer la méthode, nous montrons un résultat d'estimation de la surface du sol à partir d'un couple d'images fourni par l'ETH de Zürich. Les images scannées ont une résolution au sol voisine de 7 cm par pixel. L'image gauche du couple et le MNE issu de la reconstruction stéréoscopique sont présentés respectivement sur les figures 2.1 et 2.2.

L'estimation de la surface est réalisée à partir d'un MNE symétrisé pour éviter l'introduction de transitions altimétriques aux bords du MNE périodisé. En effet, le modèle présenté dans l'équation (2) correspond à une fonction périodisée de la surface réelle et les paramètres $b_{k,l}$ dans sont alors nuls. Nous considérons que la surface du sol peut être approximée par un modèle dont l'ordre N reste peu élevé et il convient donc de limiter l'apparition d'éventuelles hautes fréquences lors de la périodisation.

Lorsque l'ordre du modèle vaut $N = 3$ et $c_l = 1,5m$, nous obtenons le MNT représenté sur la figure 2.3. Les profils le long de la ligne A du MNE et du MNT indique que le sol est correctement estimé bien que la couverture du sur-sol sur cette scène soit proche de 40 % (voir la figure 2.4).

La méthode donne des approximations plus grossières du sol pour $N = 1,2$ mais échoue pour des ordres supérieurs à 3. En effet, pour $N > 3$ la surface paramétrique permet aussi de modéliser une partie des points du sur-sol.

Afin de modéliser des sols aux configurations variées, une alternative consiste à introduire une contrainte de régularisation. Il devient alors possible de choisir un ordre de modèle plus élevé.

4 Régularisation

Pour palier au problème de l'ordre du modèle, nous ajoutons un terme de régularisation à l'erreur globale à minimiser. L'équation (4) s'écrit alors :

$$\theta_{\rho_c, \lambda} = \arg \min_{\theta} \sum_{i \in \{1, \dots, m\}} \left\{ \rho_c(\epsilon_i, \theta) + \lambda \|\nabla z(x_i, y_i)\|_2 \right\} \quad (8)$$

où la norme euclidienne du gradient au point (x, y) est donnée par la forme analytique suivante :

$$\begin{aligned} & \|\nabla z(x, y)\|_2 \\ &= 4\pi^2 \nu_x^2 \left[\sum_{k,l=0; k+l \neq 0}^N k [-a_{k,l} S_{k,l}(x, y) + b_{k,l} C_{k,l}(x, y)] \right]^2 \\ &+ 4\pi^2 \nu_y^2 \left[\sum_{k,l=0; k+l \neq 0}^N l [-a_{k,l} S_{k,l}(x, y) + b_{k,l} C_{k,l}(x, y)] \right]^2 \end{aligned} \quad (9)$$

Le paramètre de régularisation λ dans l'équation (8) détermine l'intensité de la contrainte appliquée au modèle.

La fonction d'erreur de l'équation (8) n'étant pas linéaire, une stratégie d'optimisation alternative à la minimisation présentée dans l'équation (7) est nécessaire. À cet effet nous employons l'algorithme de Levenberg-Marquardt [8].

Nous présentons les résultats obtenus avec un couple d'images numériques. Ces images acquises lors d'un survol de la ville de Rennes par l'Institut Géographique National ont une résolution voisine de 30cm/pixel. L'image gauche

et le MNE obtenu par mise en correspondance sont présentés respectivement sur les figures 3.1 et 3.2. Dans cette scène, la différence d'altitude entre le point du sol le plus haut et le point le plus bas est proche de 5 mètres.

L'estimation robuste avec ($N = 5, \lambda = 0.5, c_l = 1,5m$) aboutit au modèle de surface présenté sur la figure 3.3. Le résultat de classification présenté sur la figure 3.4 donne une indication sur la qualité du MNT estimé. Pour réaliser la classification, un simple seuillage avec le seuil c_l est appliqué sur le MNE rectifié qui est obtenu par soustraction du MNT au MNE.

5 Conclusion

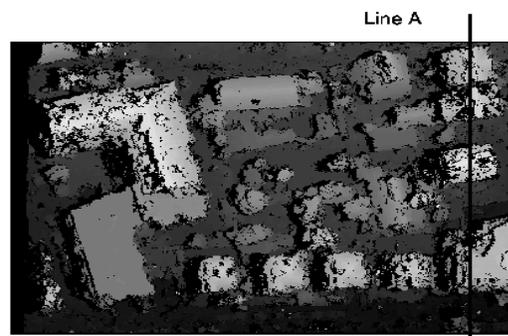
Nous avons présenté une nouvelle approche dédiée à l'extraction des objets du sur-sol à partir d'un Modèle Numérique d'Élévation. La surface du sol est modélisée par une surface paramétrique décomposée sur une base de fonctions harmoniques 2D. Une estimation robuste permet alors d'estimer les paramètres de la surface en rejetant les points du sur-sol. La méthode étant limitée vis-à-vis de l'ordre du modèle, l'introduction d'une contrainte sur le gradient de la surface rend possible la modélisation de sols plus divers. Une évaluation et une comparaison de notre méthode aux techniques mentionnées dans la section 1 est en cours.

Références

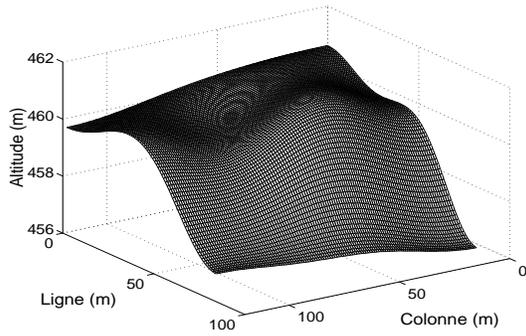
- [1] M. Cord and D. Declercq. 3d building detection and modeling using a statistical approach. *IEEE Trans. on Image Processing*, 10(5), May 2001.
- [2] U. Weidner and W. Förstner. Towards automatic building reconstruction from high resolution digital elevation models. *ISPRS Journal*, 50(4):38-49, 1995.
- [3] W. Eckstein and O. Munkelt. Extracting objects from digital terrain models. In T. Schenk, editor, *Remote Sensing and Reconstruction for Three-Dimensional Objects and Scenes*, SPIE 2572, pages 222-231, 1995.
- [4] E. Baltsavias, S. Mason, and D. Stallmann. Use of dtms/dsms and orthoimages to support building extraction. In *Automatic Extraction of Man-Made objects from Aerial and Space Images*, pages 199-210, 1995.
- [5] C. Baillard and H. Maître. 3d reconstruction of urban scenes from aerial stereo imagery: A focusing strategy. *Journal of Computer Vision and Image Understanding*, 76(3):244-258, December 1999.
- [6] T. Belli, M. Cord, and S. Philipp-Foliguet. Colour contribution for stereo image matching. In *First International Conference on Colour in Graphics and Image Processing CGIP'2000*, pages 317-322, Saint-Etienne, France, October 2000.
- [7] P.J. Huber. *Robust statistics*. John Wiley and Sons, New York, 1981.
- [8] D. Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *SIAM J. Appl. Math.*, 11:431-441, 1963.



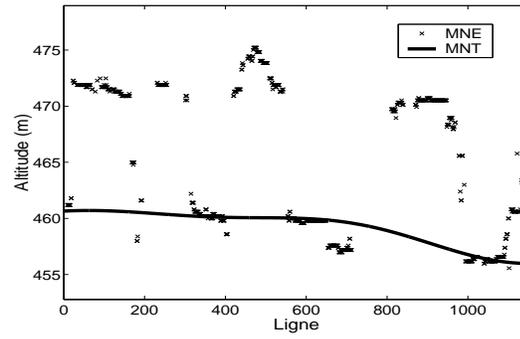
2.1: Image gauche.



2.2: MNE.



2.3: MNT.

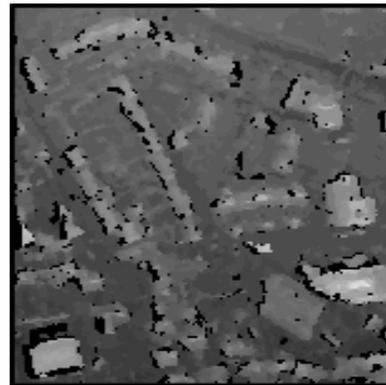


2.4: Profils selon la ligne A.

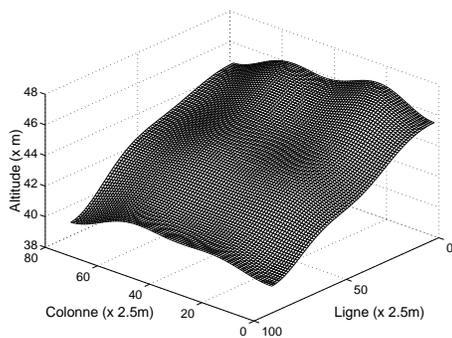
FIG. 2: Estimation du MNT à partir d'un MNE (images fournies par l'ETHZ).



3.1: Image gauche.



3.2: MNE.



3.3: MNT avec regularisation.



3.4: Classification sol/sur-sol.

FIG. 3: Estimation du MNT avec régularisation (images fournies par l'IGN).