

# Correction et amélioration des résultats de la stéréoscopie par corrélation : application au MNE urbain

Frédéric Devernay, Julien Devade

► **To cite this version:**

Frédéric Devernay, Julien Devade. Correction et amélioration des résultats de la stéréoscopie par corrélation : application au MNE urbain. Bulletin de la Société Française de Photogrammétrie et de Télédétection, SFPT, 1999, pp.52-54. <hal-00821595v2>

**HAL Id: hal-00821595**

**<https://hal.inria.fr/hal-00821595v2>**

Submitted on 14 May 2013

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Correction et amélioration des résultats de la stéréoscopie par corrélation : application au MNE urbain

Frédéric Devernay    Julien Devade

ISTAR, Les Espaces de Sophia F, 80 route des Lucioles, B.P. 282, 06905 Sophia Antipolis cedex

Phone: +33 (0)4-93-95-72-39    FAX: +33 (0)4-93-95-83-29    Email: devernay@istar.fr

## Abstract

Stereo by correlation can be used to produce urban DEM without *a priori* knowledge on the shape of buildings, as opposed to feature-based methods. Nevertheless, this method suffers from several weaknesses, e.g. at depth discontinuities located near roof borders. The classical solution is to use a correlation method with adaptive windows [Kanade et Okutomi, 1994], but we propose an alternate solution which consists in post-processing the DEM obtained by correlation. This method is based on a combination of morphomath and anisotropic diffusion techniques, and uses the local contrast of the orthoimage to determine which image areas “hooked” on the correlation window.

## Résumé

La stéréoscopie par corrélation classique permet de restituer un Modèle Numérique d'Élévation urbain (MNE) sans avoir à faire d'hypothèse a priori sur les modèles de bâtiments observés, contrairement aux méthodes à base de primitives. Cependant, cette technique souffre de plusieurs défauts, notamment au niveau des discontinuités de profondeur (par ex. au bord des toits). La solution classique consiste à utiliser une méthode de stéréoscopie à base de fenêtres adaptatives [Kanade et Okutomi, 1994], et nous proposons dans cet article une solution alternative qui consiste à faire des corrections automatiques sur le MNE *a posteriori*. Cette méthode est basée sur une combinaison de techniques de morphologie mathématique et de diffusion anisotrope, et utilise l'information de contraste local de l'orthoimage pour déterminer quels endroits ont été «accrochés» par la fenêtre de corrélation.

**Keywords:** stereoscopy by correlation, urban DEM, anisotropic diffusion, morphomath

**Mots Clés:** stéréoscopie par corrélation, MNE urbain, diffusion anisotrope, morphomath

## 1. Introduction

Pour produire des modèles numériques d'élévation urbain par stéréoscopie à partir de photographies aériennes ou d'imagerie spatiale haute résolution, il existe essentiellement deux types de méthodes. Les méthodes à base de primitives, consistent à détecter puis à apparier entre deux images ou plus des segments, des courbes, ou des objets plus complexes (par ex. des modèles de polyèdres). Les méthodes à base de stéréoscopie par corrélation consistent à mettre en correspondance des surfaces ayant des textures proches et, au contraire des méthodes à base de primitives, permettent de restituer des objets de forme quelconque, dont la surface possède des propriétés de régularité géométrique et radiométrique suffisantes. Les défauts majeurs de la stéréo par corrélation sont de mal restituer (a) les surfaces présentant des réflexions spéculaires, (b) les surfaces trop faiblement contrastées, (c) les surfaces de pente trop forte, et (d) les discontinuités de profondeur. Pour la restitution de milieux urbains à partir de photographies aériennes, ce sont surtout les défaut (b) et (d) qui dominent : les objets «bavent» hors de leur limite, notamment à cause des fortes différences de contraste entre l'arrière-plan (le sol, qui est souvent à l'ombre) et le premier plan (toits, végétation, etc.), ce qui est gênant pour certaines applications.

La solution la plus courante aux défauts (b) et (d) passe par des méthodes de corrélation utilisant des fenêtres dites «adaptatives» [Kanade et Okutomi, 1994], mais cette solution présente l'inconvénient d'être beaucoup plus complexe et moins facilement optimisable que la corrélation classique. Nous avons donc préféré une autre approche, qui consiste à utiliser la corrélation classique pour construire un MNE, puis à tenter d'améliorer a posteriori le résultat de la corrélation.

La méthode que nous avons développée utilise le fait que ce sont les zones contrastées qui «accrochent» la fenêtre de corrélation. Elle a été d'abord basée sur une méthode de diffusion anisotrope [Perona et Malik, 1990] qui utilise à la fois le MNE calculé par corrélation et une orthoimage calculée à partir de ce MNE. Nous avons ensuite ajouté des processus permettant soit d'accélérer la diffusion, soit de traiter le MNE avant la diffusion pour commencer celle-ci dans un état plus proche du résultat voulu. La méthode peut fonctionner soit sur un MNE et l'orthoimage associée, soit sur une image de disparité et l'image de référence correspondante. Les résultats montrent, sur un cas difficile (Manhattan), une nette amélioration de la qualité de restitution des contours de bâtiments et des superstructures. Le MNE résultat peut ensuite soit être utilisé tel quel, soit vectorisé manuellement ou automatiquement.

## 2. Analyse du problème

À ISTAR, les MNE urbains sont fabriqués à partir de prises de vues avec des recouvrements de 80%/80%. On produit un MNE par couple stéréo, puis on les fusionne pour former un MNE fiable (qui comporte des zones sans information). Une orthoimage est fabriquée à partir du MNE interpolé, en utilisant en chaque point l'image la plus verticale (figure 1).

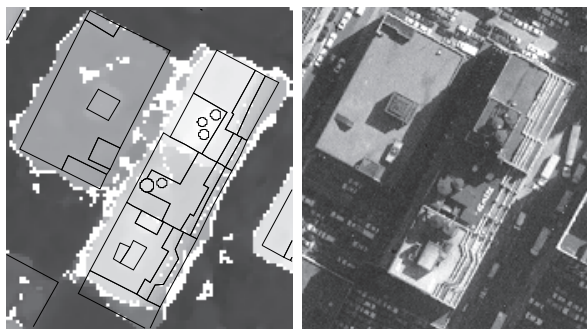


FIGURE 1 – (gauche) Le MNE brut obtenu par corrélation puis fusion de plusieurs paires stéréo de photographies aériennes : en blanc les points sans information, en rouge les contours 3D réels des bâtiments extraits manuellement à partir des mêmes images. (droite) L'orthoimage calculée à partir du MNT brut «rempli».

On remplit ensuite les zones non renseignées du MNE par interpolation contrainte par un critère de régularité, puis on calcule une orthoimage en utilisant toutes les images disponibles, pour éviter les problèmes de faces cachées (Figure 1).

Le but de la méthode proposée ici est d'améliorer la qualité du MNE ainsi calculé en utilisant à la fois le MNE et l'orthoimage. Cependant, cette méthode peut être généralisée à la situation simple où on dispose seulement d'une paire stéréoscopique : le MNE est alors remplacé par la carte de disparité, et l'orthoimage par l'image de référence de la carte de disparité.

Pour la fabrication de MNE urbain à partir de prises de vues aériennes, le principal défaut est une erreur sur le positionnement en planimétrie des points 3D. Cette erreur est due au fait que la stéréoscopie par corrélation ne donne pas la disparité du point situé au centre de la fenêtre : si une zone de la fenêtre de corrélation est plus texturée ou plus contrastée que le reste, c'est plutôt la disparité de cette partie-là qu'on obtiendra. On peut donc avoir une erreur d'estimation planimétrique allant jusqu'à la moitié de la taille de la fenêtre de corrélation, dans les directions  $x$  et  $y$  de l'image (qui correspondent aux axes cartographiques  $X$  et  $Y$  dans notre cas).

Ce défaut se manifeste plus particulièrement aux contours des bâtiments. En effet, ce sont des discontinuités de profondeur, et les deux côtés de la discontinuité (sol et toit) on souvent des contrastes très

différents, dus à la différence d'éclairage (ombres portées au sol)). Le plus souvent, donc, les bâtiments vont «baver», surtout du côté des ombres, et leur taille sera plus large que la taille réelle.

Un autre défaut souvent observé est que dans les zones d'ombres, où la dynamique est faible, la corrélation ne donne aucun résultat. Si la rue est étroite ou que les bâtiments sont hauts, on peut n'avoir presque aucune information au niveau de la rue (en haut à droite de la Figure 1).

La solution la plus courante consiste à utiliser les fenêtres adaptatives [Kanade et Okutomi, 1994]. Dans cette méthode, la taille de la fenêtre de corrélation diminue près des contours des images, et augmente aux endroits de faible contraste. Cette méthode fonctionne bien en milieu urbain, mais est plutôt complexe à mettre en œuvre, ce qui nous a poussé à étudier une autre solution.

Cette solution est basée sur l'observation suivante : les zones que la fenêtre de corrélation a «accroché» correspondent aux zones fortement contrastées de l'orthoimage. On doit donc pouvoir chercher à corriger le MNE, en utilisant l'orthoimage, en modifiant les valeurs de disparité qui se situent dans des zones peu contrastées proches de zones contrastées. Ces zones à corriger sont de largeur au plus égale à la moitié de la taille de la fenêtre de corrélation rapportée au sol.

## 3. Approche générale

La première approche est basée sur une technique de diffusion anisotrope [Perona et Malik, 1990], qui agit sur le MNE, et dont le coefficient de conduction est calculé à partir de l'orthoimage. Cette approche est générale et peut s'appliquer à tout type de MNE (même non urbain), voire même à une paire stéréoscopique quelconque avec la carte de disparité associée.

On commence par filtrer le MNE brut, contenant déjà des zones non renseignées dues aux erreurs de corrélation, pour enlever les valeurs aberrantes, par l'élimination des valeurs trop peu représentées dans un histogramme local du MNE. On remplit ensuite les zones non renseignées en tenant compte du contraste local de l'orthoimage, puis on applique ensuite un nombre fini d'itérations de diffusion anisotrope.

Le remplissage et la diffusion sont basés sur une image représentant le contraste local de l'image de référence (l'orthoimage). Cette mesure doit avoir une précision de localisation meilleure que la taille de la fenêtre de corrélation, puisqu'on cherche à déterminer quels sont les endroits qui ont pu «accrocher» la corrélation à l'intérieur-même de cette fenêtre. Plusieurs essais infructueux en milieu urbain avec diverses mesures de contraste nous ont amené à utiliser un simple détecteur de contours, le filtre de Kisch.

Lors des opérations de remplissage et de diffusion, un seuil est utilisé ( $\kappa$ , le même pour les deux opérations)

qui correspond à la limite dans l'image de contraste entre les zones contrastées et les zones peu contrastées (ombres).

La deuxième phase consiste à remplir les valeurs «invalides», dues à des erreurs de corrélation ou résultant du filtrage, par des valeurs probables. Ce remplissage consiste en une succession de dilatations basées sur l'image de contraste : un pixel du MNE dont le contraste est inférieur à un seuil  $\kappa$  est rempli par la médiane de ses voisins valides dont le contraste est inférieur à  $\kappa$ . Le seuil  $\kappa$  est progressivement augmenté, jusqu'à ce que tout le MNE soit rempli de valeurs valides.

Cette méthode de remplissage a pour effet de remplir les zones invalides de préférence à partir des zones de contrastes équivalent. En effet, un pixel situé à côté d'un pixel plus contrasté et valide et qui est effectivement à la même hauteur a peu de chance d'être invalide. Il est donc raisonnable de faire hériter aux pixels invalides l'altitude de pixels moins contrastés. Les pixels invalides qui correspondent à des maxima locaux de contraste ne peuvent être remplis par ce processus, et on leur affecte l'altitude médiane de leurs voisins valides. L'étape suivante consiste à modifier légèrement ces valeurs valides pour tenir compte des effets dus à la taille de la fenêtre de corrélation, par diffusion.

La diffusion consiste à observer l'évolution dans le temps d'une grandeur scalaire (en général l'intensité) définie dans un espace de dimension  $n$  (2 pour une image), soumise à une loi d'évolution similaire à l'équation de la chaleur. La diffusion anisotrope [Perona et Malik, 1990] consiste à ajouter un coefficient de conduction calculé à partir des dérivées de l'image, de manière à ce que la diffusion conserve les contours présents dans l'image initiale.

Dans notre cas, nous avons à notre disposition deux grandeurs scalaires : l'altitude du MNE et l'intensité de l'orthoimage. Nous allons faire diffuser l'image d'altitude, en utilisant l'image d'intensité (plus précisément de contraste) pour calculer le coefficient de conduction. Ainsi, à un endroit où l'image est très contrastée, la donnée d'altitude est en général fiable, et elle ne doit pas être modifiée. Par contre, si il y a à une distance inférieure à la taille de la fenêtre de corrélation une zone peu contrastée, elle aura probablement été masquée par la zone contrastée lors de la stéréoscopie par corrélation, et son élévation doit donc être corrigée à partir de l'élévation des zones peu contrastées de son voisinage.

En appliquant cette formule de diffusion, on note de nettes améliorations du MNE, notamment au niveau des contours des bâtiments, des superstructures, et de la végétation qui sont beaucoup plus francs. Cependant, la méthode et les résultats souffrent de plusieurs défauts en milieu urbain : le nombre d'itérations nécessaires pour résorber les erreurs dues aux immeubles les plus hauts est important (plus de 200) ; à cause du grand nombre d'itérations, les er-

reurs du MNE ont une zone d'influence dans l'image qui va bien au-delà de la taille de la fenêtre de corrélation ; si un immeuble n'a pas un contour partout contrasté, l'élévation «fuit» par les ouvertures et son altitude diminue ; de même, les débordements des immeubles peuvent, par diffusion, relever le niveau du sol de plusieurs mètres. Ces défauts disparaissent lorsqu'on applique la méthode à un MNE non urbain ou de plus faible résolution (fait par exemple à partir d'images SPOT).

#### 4. Améliorations, résultats et conclusion

En milieu urbain, les zones peu contrastées sont souvent des ombres, et sont donc au-dessous des zones contrastées. La première amélioration de la méthode consiste donc à éliminer une partie des bavures de la corrélation par segmentation et érosion. Le MNE est d'abord segmenté en zones de contraste inférieur au seuil  $\kappa$  et homogènes. On fait ensuite quelques érosions sur toutes les régions qui ont une voisine de cardinal supérieur et d'altitude inférieure (ces régions sont considérées comme des bavures).

La seconde amélioration (méthode en deux passes) consiste à éliminer les points du MNE dont la variation entre l'initialisation et le résultat final a été importante, puis à réitérer le processus.

La figure 2 montre le résultat de la méthode sur un bâtiment de Manhattan en une et deux passes. La méthode a été appliquée à l'ensemble de la presqu'île, à une résolution de 50cm, et a montré qu'un simple post-traitement pouvait améliorer sensiblement les résultats de la stéréoscopie par corrélation classique.

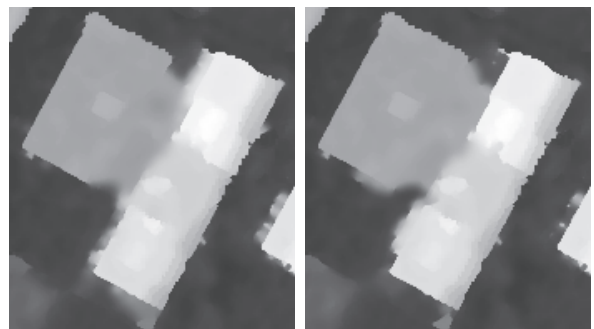


FIGURE 2 – Le MNE après traitement : (gauche) filtrage, remplissage, régionnement et érosions, diffusion. (droite) en utilisant la méthode en deux passes.

#### Références

- [Kanade et Okutomi, 1994] T. Kanade et M. Okutomi. A stereo matching algorithm with an adaptive window : Theory and experiment. *IEEE Trans. PAMI*, 16(9) :920–932, Sept. 1994.
- [Perona et Malik, 1990] P. Perona et J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Tr. PAMI*, 12(7) :629–639, July 1990.