
La morphologie mathématique binaire pour l'extraction automatique des bâtiments dans les images THRS

David Sheeren^{*,**} – Sébastien Lefèvre^{*} – Jonathan Weber^{*}

**Laboratoire des Sciences de l'Image, de l'Informatique et de la télédétection
UMR 7005 CNRS – Université Louis Pasteur, Strasbourg I.
Pôle API, Boulevard S. Brant, BP 10413 – 67412 Illkirch Cedex
lefevre@lsiit.u-strasbg.fr*

***UMR 1201 Dynamiques Forestières dans l'Espace Rural, INRA – INPT/ENSAT
Av. de l'Agrobiopôle, BP 32607, Auzeville Tolosane, 31326 Castanet Tolosan cedex
David.Sheeren@ensat.fr*

RÉSUMÉ. Cet article présente une nouvelle méthode de détection et d'extraction des bâtiments en milieu urbain à partir d'images satellitaires à très haute résolution spatiale. L'approche proposée est fondée sur l'application et l'enchaînement automatique d'opérateurs issus de la morphologie mathématique binaire. Plusieurs étapes constituent la méthode : (1) binarisation de l'image, (2) filtrage du bruit et des éléments de taille inférieure aux bâtiments par lissage morphologique, (3) détection des bâtiments par application d'une transformée en tout ou rien adaptative, avec un élément structurant de taille et de forme variable, (4) restauration de la forme des bâtiments par reconstruction géodésique. Deux stratégies différentes de binarisation sont proposées lors de l'étape initiale. La première consiste à binariser l'image par seuillage, le seuil étant défini soit de manière automatique, soit de manière empirique en fonction de l'image traitée. La seconde est fondée sur l'application d'une classification non supervisée pour laquelle le nombre de classes n'est pas fixé a priori. La méthode a été mise en œuvre sur une image Quickbird panchromatique de la région de Strasbourg. Les résultats obtenus confirment l'intérêt et l'efficacité de l'approche.

ABSTRACT. This paper presents a new method for building extraction in Very High Resolution remotely sensed images in urban areas. The approach proposed is based on the use binary mathematical morphology operators. The method is composed of several steps: (1) conversion of grey level images to binary images, (2) smoothing by means of morphological filtering, (3) building detection with an adaptive hit-or-miss transform, (4) shape restoration. Two strategies of binarization are proposed. The first one consists in performing an interactive or automatic thresholding. The second one is based on an unsupervised classification. The method has been applied on a Quickbird panchromatic image. Results show the interest of the approach.

MOTS-CLÉS : télédétection, morphologie mathématique, transformée en « tout ou rien ».

KEYWORDS: remote sensing, mathematical morphology, hit-or-miss transform.

1. Introduction

L'intérêt d'utiliser la télédétection aérospatiale pour étudier l'espace urbain n'est plus à démontrer, même si les applications dans ce milieu restent relativement récentes [Weber 1995, Donnay *et al.* 2001]. Ainsi, outre l'identification des éléments du tissu urbain, elle peut servir à cartographier les changements d'utilisation du sol, par analyses d'images multi-temporelles, et donc aider à prévoir et planifier la croissance urbaine. La morphologie des villes, leurs structures internes et leurs limites peuvent également être détectées et analysées. Dans le contexte actuel de développement durable, la télédétection offre des moyens aux décideurs de gérer et d'aménager leur territoire d'une manière plus soucieuse de l'environnement.

Depuis l'apparition des satellites à très haute résolution spatiale (THRS) tels que SPOT5, Quickbird ou Ikonos, la palette d'applications de la télédétection en milieu urbain s'est sensiblement élargie. Parallèlement, les méthodes d'analyse et traitement des images ont aussi évoluées. Les entités géographiques n'étant plus représentées que par un seul pixel, les approches dites « orientée objet » ont été proposées, où chaque objet géographique se compose d'un agrégat de pixels homogènes [Benz *et al.* 2004]. La reconnaissance de ces objets n'est plus seulement fondée sur leurs propriétés spectrales (signature, texture...) comme c'est le cas dans les méthodes traditionnelles. Les caractéristiques spatiales et contextuelles des objets sont également prises en compte (forme, taille, relation topologique, proximité...). Toutefois, pour pouvoir exploiter cette variété de critères efficacement et ainsi améliorer les résultats de classification, il est nécessaire d'exploiter des images bien segmentées. Or, s'il existe aujourd'hui de nombreuses méthodes de segmentation différentes [Cocquerez et Philipp 1995], la qualité des images segmentées n'est pas encore complètement satisfaisante [Neubert *et al.* 2006]. Les entités géographiques sont rarement correctement individualisées. Une entité peut ainsi être morcelée en plusieurs éléments composant différents segments. Plusieurs entités peuvent également être fusionnées en un seul segment.

Plutôt que d'adopter une approche de segmentation globale plus générique, mais générant des défauts dans l'extraction des différentes catégories d'objets, il est possible d'envisager d'utiliser des méthodes plus spécifiques, adaptées seulement à un certain type d'objets, mais fournissant des résultats de meilleure qualité. Les travaux de recherche que nous présentons dans cet article s'inscrivent dans ce contexte.

Nous proposons une approche qui permet d'extraire automatiquement les bâtiments dans des images satellitaires THRS en milieu périurbain, en nous focalisant en particulier sur les bâtiments pavillonnaires. Cette approche est fondée sur l'utilisation d'opérateurs issus de la morphologie mathématique binaire. Elle ne requiert aucune donnée exogène en dehors de l'image originale et est compatible avec une large variété d'images satellitaires.

L'article est structuré de la manière suivante. La section 2 est consacrée au rappel des principes d'analyse d'images par la morphologie mathématique dans laquelle nous mentionnons plusieurs travaux antérieurs en télédétection qui exploitent ces outils. Nous présentons en section 3 l'approche proposée qui se caractérise par un enchaînement de différentes étapes : (1) binarisation de l'image, (2) lissage morphologique, (3) application d'une transformée en « tout ou rien » adaptative, (4) reconstruction géodésique. La section 4 illustre l'application de la méthode et les principaux résultats obtenus sur une image panchromatique (Quickbird). Nous concluons l'article en suggérant plusieurs perspectives de recherche en section 5.

2. Morphologie mathématique

L'analyse d'images par la morphologie mathématique remonte aux années soixante. Cette théorie a été initialement introduite en sciences des matériaux par J. Serra et G. Matheron pour analyser des objets au travers de leur texture (Serra 1982). Elle a connu durant ces deux dernières décennies de nombreux développements, tant sur le plan théorique que pratique. Elle couvre aujourd'hui un éventail de champs d'application assez large, touchant notamment à la robotique et la vision industrielle, l'imagerie médicale ou encore le multimédia [Soille, 2004].

Comme en témoigne les nombreux travaux publiés, la télédétection aérospatiale est également devenu un domaine d'application privilégié [Destival 1986, Flouzat 1988, Soille et Pesaresi 2002, Benediktsson *et al.* 2003]. L'intérêt de la morphologie mathématique dans ce domaine n'a rien d'étonnant puisqu'elle offre la possibilité de rechercher dans une image des configurations spatiales prédéfinies. Elle est donc particulièrement adaptée à l'extraction d'entités dont les formes ont une signification thématique. C'est par exemple le cas des routes [Zhang *et al.* 1999, Katartzis *et al.* 2000], des traits de côtes [Jishuang et Chao 2002], des bâtiments [Brunn et Weidner 1997, Matti-Gallice et Collet 2004, Jin et Davis 2005] ou encore de différentes structures paysagères [Callot *et al.* 1993, Zheng *et al.* 1995, Laporterie *et al.* 2001].

Les méthodes de détection du bâti existantes requièrent souvent des données spécifiques, comme des modèles numériques de terrain [Brunn et Weidner, 1997] ou une intervention forte de l'utilisateur. La méthode proposée par [Matti-Gallice et Collet 2004] nécessite ainsi de réaliser une classification supervisée pour affiner les résultats fournis par les opérateurs morphologiques. Certaines approches sont également fondées sur des hypothèses parfois contraignantes, comme celle de [Jin et Davis, 2005] où les bâtiments sont partiellement identifiés grâce à leur brillance et leur ombre. La méthode que nous proposons ici utilise également des outils relevant de la morphologie mathématique, mais diffère des approches existantes en ce sens qu'elle ne nécessite aucune information externe et s'applique sur des images THRS.

2.1. Principes

L'analyse d'une image par morphologie mathématique consiste à appliquer sur cette image un ou plusieurs opérateurs morphologiques au moyen d'un masque ou motif géométrique prédéfini (carré, rectangle, cercle...). Le motif géométrique, appelé *élément structurant* (ES), est déplacé dans l'image de façon à ce que son centre occupe toutes les positions de l'espace (les pixels p). Pour chacune des positions, on applique l'opérateur choisi. Il en résulte une image transformée dans laquelle des éléments X peuvent avoir été éliminés, fusionnés, amincis, etc. (figure 1).

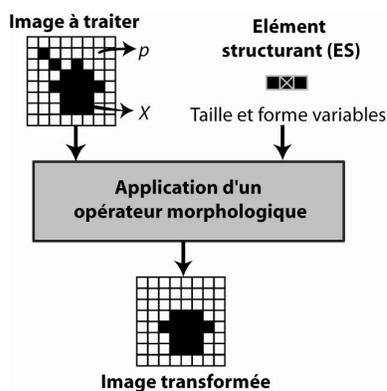


Figure 1. Principe de l'analyse d'images par la morphologie mathématique.

On distingue la morphologie mathématique ensembliste qui porte sur des images binaires, de la morphologie mathématique fonctionnelle qui opère sur des images en niveaux de gris. La première fait appel à la théorie des ensembles. Pour chaque position du centre de l'élément structurant, on vérifie une relation ensembliste (union, intersection) de l'ES avec les éléments de l'image. Le résultat de la transformation est une nouvelle image binaire composée des points qui satisfont la relation. Pour les images en niveaux de gris, les transformations morphologiques sont généralisées et s'appliquent sur des fonctions (et non plus des ensembles) en tenant compte cette fois des valeurs prises par les pixels dans le domaine de l'élément structurant. Suivant l'opérateur choisi, on affectera par exemple la valeur maximale des pixels au centre de l'ES (cas d'une dilatation plate) ou la valeur minimale (cas d'une érosion plate).

Puisque nous cherchons ici à montrer l'intérêt de la morphologie mathématique pour détecter des bâtiments en s'appuyant uniquement sur des informations de forme, la méthode que nous proposons ne fait appel qu'à des opérateurs morphologiques binaires. En effet, l'utilisation d'opérateurs morphologiques en niveaux de gris fournit également des informations relatives à la texture des objets car elle intègre

des statistiques d'ordre (minimum, maximum) calculées localement. Cette information de texture peut évidemment être pertinente pour la détection du bâti, mais son exploitation n'est pas envisagée ici. La suite de l'exposé s'inscrit donc dans le cadre de la morphologie mathématique binaire.

2.2. Opérateurs morphologiques fondamentaux

Les deux opérateurs élémentaires de la morphologie mathématique sont l'*érosion* et la *dilatation*. Ils sont notés respectivement $A \ominus B$ et $A \oplus B$ où A correspond à l'image binaire à traiter et B l'élément structurant avec lequel on parcourt l'image. Ces opérateurs sont définis par les équations suivantes :

$$\begin{aligned} A \ominus B &= \{x : (B)_x \subseteq A\} \\ A \oplus B &= \{x : (B')_x \cap A \neq \emptyset\} \end{aligned}$$

où $(B)_x$ représente B translaté par x , et B' la réflexion (symétrie centrale) de B .

L'érosion a pour effet d'éliminer les objets ou certaines parties d'objets dont la taille est inférieure à celle de l'élément structurant (suite à la relation d'inclusion totale à vérifier). Certains objets connectés peuvent également être séparés. À l'inverse, la dilatation a pour effet d'augmenter la taille des objets, et de fusionner certains d'entre eux, lorsque la distance qui les sépare est inférieure à la taille de l'élément structurant (intersection non vide).

À partir de ces opérateurs, une série d'autres opérateurs sont définis comme des combinaisons plus ou moins complexes d'érosions et de dilatations. Ainsi, l'*ouverture* morphologique, qui vise à supprimer les détails de taille inférieure à l'ES (sans pour autant réduire la taille des autres éléments) correspond à une érosion suivie d'une dilatation :

$$A \circ B = (A \ominus B) \oplus B$$

L'opération duale, qui correspond à la *fermeture* morphologique, se définit comme une combinaison d'une dilatation suivie d'une érosion :

$$A \cdot B = (A \oplus B) \ominus B$$

Il faut noter que l'ouverture et la fermeture sont idempotentes autrement dit, le résultat est invariant après transformations successives.

La figure 2 illustre l'effet des différents opérateurs morphologiques définis ci-dessus (l'image originale étant l'image à traiter de la figure 1).

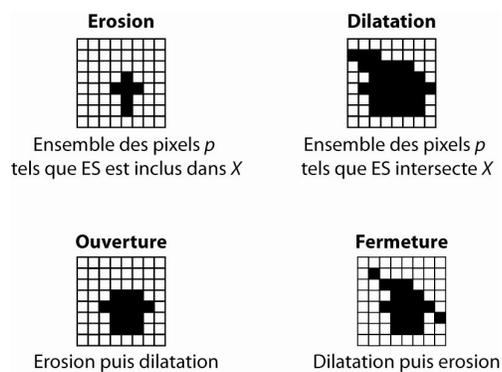


Figure 2. Résultats d'opérations morphologiques élémentaires.

2.3. Transformée en « tout ou rien »

La méthode que nous proposons pour extraire automatiquement les bâtiments (section 3) fait intervenir une adaptation d'un opérateur classique de la morphologie mathématique : la transformée en « tout ou rien » (TTR) [Serra 1982]. Cette transformée consiste en l'application d'une double érosion (sur l'image A et son complémentaire A^c) avec deux éléments structurants disjoints ayant la même origine, notés E et F :

$$A \otimes (E, F) = (A \ominus E) \cap (A^c \ominus F)$$

L'opérateur cherche donc à vérifier que les éléments de A sont inclus dans E et que les éléments de A^c (le fond de l'image) sont inclus dans F . L'ensemble des positions pour lesquelles ces relations sont respectées participent à la définition de l'image résultante (figure 3).

La TTR est souvent appliquée en reconnaissance de formes pour détecter des motifs de forme prédéfinie. La difficulté consiste naturellement à déterminer l'élément structurant puisque la forme des objets peut varier et que le processus doit être robuste au bruit.

Notre méthode d'extraction du bâti est fondée sur cet opérateur morphologique qui semble n'avoir jamais été utilisé auparavant dans le domaine de l'imagerie spatiale pour résoudre le problème étudié. Nous exposons l'approche proposée ci-dessous.

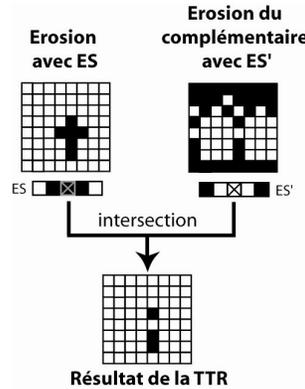


Figure 3. La transformée en « tout ou rien » avec deux éléments structurants ES et ES' disjoints.

3. Méthode d'extraction automatique du bâti

La méthode proposée est destinée principalement aux images satellitaires à très haute résolution spatiale et vise à détecter les bâtiments individualisés en milieu urbain et périurbain. Elle n'est pas encore adaptée à l'extraction de zones d'espace bâti en milieu urbain dense, pour lesquelles il est difficile de déterminer a priori des éléments structurants de forme et de taille pertinents.

Notre méthode se caractérise par un enchaînement d'étapes, chacune faisant appel à différents opérateurs morphologiques. Les différentes étapes sont illustrées en figure 4 et détaillées dans les sections suivantes.

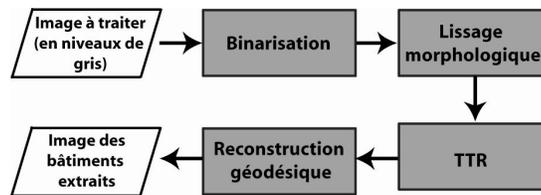


Figure 4. Démarche de la méthode d'extraction proposée.

3.1. Etape 1 : binarisation

Puisque nous exploitons des opérateurs de la morphologie mathématique binaire, l'image I fournie en entrée doit être binarisée. Cette opération constitue la première étape de la méthode. Nous proposons deux stratégies différentes pour y procéder. La

première est une binarisation classique, par seuillage. La seconde est une binarisation fondée sur une classification non supervisée de l'histogramme de l'image. Dans les deux cas, nous appliquons la binarisation sur une image filtrée par un lissage morphologique. Celui-ci est défini par l'opérateur OCCO (*open-close, close-open*) qui combine des ouvertures et des fermetures de la façon suivante :

$$\text{OCCO}_B(A) = \left[\frac{1}{2}((A \cdot B) \circ B) + \frac{1}{2}((A \circ B) \cdot B) \right]$$

3.1.1. Binarisation par seuillage

La première stratégie consiste à binariser l'image par seuillage et à garder les pixels dont l'intensité est supérieure au seuil fixé. On regroupe donc au sein d'une même classe, des pixels de valeurs comprises dans un intervalle donné de l'histogramme. Le seuil est généralement défini par l'utilisateur en tenant compte des caractéristiques de l'image et en s'assurant que les objets d'intérêts (les bâtiments) sont pris en compte.

Cette stratégie de binarisation présente l'avantage d'être simple et rapide à mettre en œuvre. On peut directement décider de l'appartenance d'un pixel au fond ou à l'objet. Cependant, le choix du seuil est une opération délicate. Des parties d'objets d'intérêts peuvent être affectées au fond de l'image si les pixels ont une intensité légèrement inférieure au seuil fixé. Par ailleurs, ce choix n'est généralement pas automatique. Il est laissé à l'utilisateur qui le détermine de manière empirique. Puisque le seuil peut varier d'une image à l'autre (en fonction du type de capteur, de la zone étudiée, de l'ensoleillement...) l'utilisateur doit examiner chaque nouvelle image avant de procéder au seuillage, ce qui est naturellement coûteux en temps. Il est également possible d'effectuer un seuillage automatique (à l'aide de techniques bien connues comme [Ridler et Calvard, 1978] ou [Otsu, 1979]). Cependant, même si le processus gagne en généralité, l'image binaire obtenue possède les mêmes défauts qu'avec un seuillage manuel. Cette méthode de binarisation est en fait une méthode globale. Elle suppose que seulement deux classes spectrales sont représentées dans l'image, et que les objets peuvent être identifiés par une signature spectrale unique. Cette approche n'est pas bien adaptée à des images satellitaires THRS qui peuvent présenter une hétérogénéité spectrale importante et pour lesquelles un simple seuillage (qu'il soit automatique ou empirique) ne permet pas de dissocier correctement les pixels représentant le bâti des autres pixels [Weber et al. 2006].

La deuxième stratégie de binarisation proposée a pour objectif de pallier à ces différents inconvénients. Elle est automatique, plus générique, et exempt de seuil. Nous la présentons ci-dessous.

3.1.2. Binarisation par classification non supervisée d'histogramme

Cette méthode de binarisation est fondée sur la combinaison d'un ensemble d'images binaires différentes, résultant d'une classification non supervisée de l'image initiale (en niveaux de gris) établie à partir de l'analyse automatique de son histogramme.

Le principe de la méthode est le suivant :

- l'image lissée initialement par l'opérateur OCCO est d'abord classée. La classification, qui est non supervisée, est réalisée à partir de l'histogramme de l'image. Au préalable, et afin d'accroître la robustesse au bruit de la méthode, l'histogramme lui-même est également lissé à l'aide d'opérations morphologiques, ce qui permet de supprimer les maxima ou minima locaux non significatifs. Les classes (*clusters*) dont le nombre n'est pas fixé a priori sont ensuite identifiées de manière itérative, en sélectionnant les modes de l'histogramme, ou maxima locaux les plus importants. Chaque maximum local accompagné de ses valeurs voisines est affecté à un cluster et est ensuite annulé dans l'histogramme (c'est-à-dire mis à 0). Lorsqu'il ne reste plus qu'un certain pourcentage de pixels à classer, la procédure itérative s'arrête et les pixels restants sont affectés aux clusters les plus proches. On obtient ainsi un ensemble d'images binaires (une par cluster) qui représentent l'appartenance ou non des pixels aux clusters.
- Les clusters extraits peuvent ne pas représenter à eux seuls les bâtiments. Des parties de ceux-ci peuvent se retrouver dans plusieurs clusters en raison des différences de réflectance de leurs éléments constitutifs. On ajoute donc à l'ensemble des images binaires (une par cluster) un autre ensemble d'images obtenues par fusion. Plus précisément, on considère toutes les combinaisons (de cardinalité maximale prédéfinie) possibles d'images binaires, et pour chaque combinaison, on procède à une fusion des images binaires (appelée fusion de clusters) à l'aide d'une union. Un pixel est donc blanc dans l'image fusionnée si et seulement si il l'est dans au moins une image de la combinaison considérée. On obtient au final un ensemble d'images binaires plus complexes avec lesquelles on poursuit individuellement le processus d'extraction des bâtiments (cf. étapes suivantes)

Cette méthode de binarisation présente plusieurs avantages par rapport à la précédente. D'abord, elle est automatique et ne requiert pas de seuil global pour être mise en œuvre. Le paramètre à fixer par l'utilisateur est le pourcentage de pixels qui doivent être affectés aux clusters pour que la procédure itérative s'arrête. Ce paramètre peut être défini par défaut et convenir pour différentes images. Il est également beaucoup plus robuste qu'un simple seuil. Ensuite, cette méthode de binarisation prend mieux en compte l'hétérogénéité des valeurs spectrales des éléments constitutifs des bâtiments (et notamment des toits). La fusion des clusters permet d'obtenir une image binaire dans laquelle figure toutes les portions d'objets

qui composent les bâtiments, ceux-ci pouvant être caractérisés par des signatures spectrales très différentes.

3.2. Etape 2 : Lissage morphologique

L'image étant binarisée, celle-ci est ensuite filtrée avant de procéder à la détection proprement dite des bâtiments. C'est la deuxième étape de la méthode. Ce filtrage vise à éliminer les objets dont la taille est inférieure à la taille minimum supposée d'un bâtiment dans la scène étudiée. Il est effectué en appliquant une ouverture morphologique avec un élément structurant de forme carrée. La taille de cet élément structurant est déterminée par l'utilisateur, en fonction des caractéristiques de l'image binarisée.

Dans le cas où la binarisation est effectuée par classification non supervisée d'histogramme, chacune des images obtenues suite à la fusion de clusters fait l'objet du lissage décrit ci-dessus.

3.3. Etape 3 : Extraction des bâtiments avec la TTR

C'est à partir de cette image filtrée que se déroule la procédure d'extraction des bâtiments proprement dite. Il s'agit de la troisième étape de la méthode. Nous proposons d'appliquer une TTR en prenant en compte des ES de tailles et formes variables. Cette adaptation de la TTR classique (pour laquelle l'ES est figé) permet ainsi de détecter une grande variété de bâtiments.

Ainsi, on définit les ES $E_{a,b}$ et $F_{c,d}$ respectivement comme un rectangle de taille $a \times b$ et comme un cadre (contour d'un rectangle) de taille $c \times d$ avec $c > a$ et $d > b$. Le résultat de l'application de la TTR à ES variables est alors défini comme l'union des différentes applications de la TTR, soit :

$$A \otimes_{\mathcal{K}, \mathcal{L}} (E, F) = \bigcup_{\substack{k \in \mathcal{K} \\ l \in \mathcal{L}}} (A \ominus E_{\alpha k, \alpha l}) \cap (A^c \ominus F_{k, l})$$

où $K \times L$ est l'ensemble des tailles considérées pour E et F (K et L représentant respectivement les hauteurs et les largeurs), et α un coefficient utilisé dans la définition de la zone incertaine entre E et F , zone non considérée par la TTR.

A l'aide de cette transformée, les régions qui sont conservées sont celles qui peuvent contenir l'ES E et qui sont encadrées par l'ES F . La zone incertaine entre les ES E et F garantit au processus une certaine tolérance dans le contenu ou la forme des régions analysées.

Comme lors de l'étape précédente, le traitement décrit ici est appliqué sur chaque image binaire dans le cas d'une binarisation par classification et fusion des clusters.

3.4. Etape 4 : Reconstruction de la forme des bâtiments

L'application de la TTR permet de supprimer les régions de l'image qui ne sont pas des bâtiments et de ne garder que les objets d'intérêt. Toutefois, elle a aussi pour effet de modifier la forme originale des bâtiments existants puisqu'elle est fondée sur des érosions qui ont pour effet de réduire la taille des objets présents dans l'image. L'objectif de cette quatrième étape est donc de restaurer la forme originale des bâtiments.

L'opération est réalisée en appliquant une *reconstruction géodésique* qui repose sur l'utilisation de deux images : une image masque et une image marqueur. L'opération consiste à effectuer une succession de dilations sur l'image marqueur et de combiner le résultat de ces dilations avec le masque. Le résultat de la reconstruction géodésique correspond à l'image marqueur dilatée et conditionnée par le masque.

Pour la reconstruction des bâtiments, l'image marqueur est composée uniquement des éléments à retenir mais dont la forme a été modifiée par l'étape 3, tandis que l'image masque contient les éléments non déformés mais aussi d'autres éléments à éliminer, c'est-à-dire le résultat fourni par l'étape 2. Plus formellement, l'opération de reconstruction géodésique est appliquée ici sur l'image filtrée par l'ouverture lors de l'étape 2 (masque), en considérant comme marqueur le résultat de la TTR :

$$C \triangle_B A = (A \oplus_C B)^\infty$$

avec une application jusqu'à idempotence (invariance de la transformation) de la dilatation conditionnelle définie par :

$$A \oplus_C B = (A \oplus B) \cap C$$

Nous obtenons ainsi tous les bâtiments détectés lors de la TTR avec une forme correspondant à celle obtenue à l'issue du lissage morphologique.

Dans le cas où la binarisation a été effectuée par classification d'histogramme, on dispose de plusieurs résultats, un par image binaire générée initialement par la fusion de clusters. Une étape supplémentaire d'unification des résultats est alors considérée : le résultat final est obtenu par l'union des résultats intermédiaires. On conserve donc tous les objets détectés dans l'une ou l'autre des images binaires après la reconstruction géodésique.

4. Applications

Afin d'évaluer la méthode définie, celle-ci a été appliquée sur une image THRS. Nous présentons dans cette partie les principaux résultats obtenus.

4.1. Données

La reconnaissance automatique de bâtiments pavillonnaires, pour qu'elle soit efficace, requiert l'exploitation d'images THRS (< 5m) étant donné que ces objets ne présentent pas une grande étendue. C'est la raison pour laquelle nous avons utilisé des données provenant d'un capteur THRS, Quickbird PAN, offrant une résolution spatiale de 0,7m. Bien que le satellite Quickbird fournisse également des données multispectrales, nous nous sommes limités ici à n'utiliser que l'image panchromatique. C'est effectivement cette dernière qui présente la plus grande résolution spatiale et qui contient l'information géométrique. Elle est donc mieux adaptée à la reconnaissance de formes à l'aide d'opérateurs morphologiques.

Les zones d'étude sur lesquelles ont porté les expérimentations se situent dans la région urbaine de Strasbourg (France). Les sites choisis ont été sélectionnés en fonction leur complexité et en s'assurant d'avoir une diversité représentative du milieu urbain. Certains sites ont une composition relativement homogène. D'autres sont plus hétérogènes avec des bâtiments de tailles et formes variables, composés de toits en matériaux différents. Une illustration de quelques extraits de données utilisés pour les tests est fournie en figure 5. Le jeu de données contient au total 84 bâtiments.

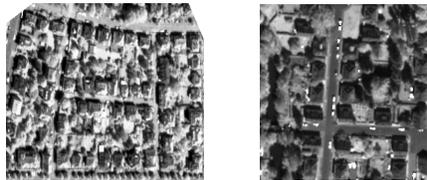


Figure 5. Extraits de quelques jeux de données utilisés pour l'application de la méthode proposée.

4.2. Mise en œuvre de la méthode et principaux résultats

Un examen visuel des données a d'abord été mené avant d'appliquer la méthode d'extraction des bâtiments. Cet examen fut nécessaire pour fixer les différents paramètres requis par les opérateurs morphologiques utilisés (taille et forme des ES), lesquels sont notamment fonction de la taille du plus petit bâtiment présent dans

l'image traitée. Nous avons fait ainsi le choix de n'appliquer que des ES de forme carrée et rectangulaire de tailles variables (par application de la TTR adaptative), sur des bâtiments orientés horizontalement ou verticalement. Nous nous sommes volontairement restreints dans cette étude à ce type de motif afin de simplifier les expérimentations et valider l'approche dans des conditions bien définies. Les données testées ont été choisies en respectant ce critère. Une rotation des images a donc été appliquée quand ce fut nécessaire.

La figure 6 illustre l'application de la méthode en suivant la première stratégie de binarisation, par seuillage. La taille des ES relatifs au lissage par ouverture et à la détection par TTR est fonction de la taille minimale et maximale des bâtiments présents dans l'image considérée, qui sont deux mesures à régler empiriquement par analyse visuelle de l'extrait traité. A titre d'exemple, la taille de l'ES choisi pour détecter les bâtiments de la figure 6 varie de 15x15 à 30x30 pixels (la longueur et la largeur évoluant indépendamment).

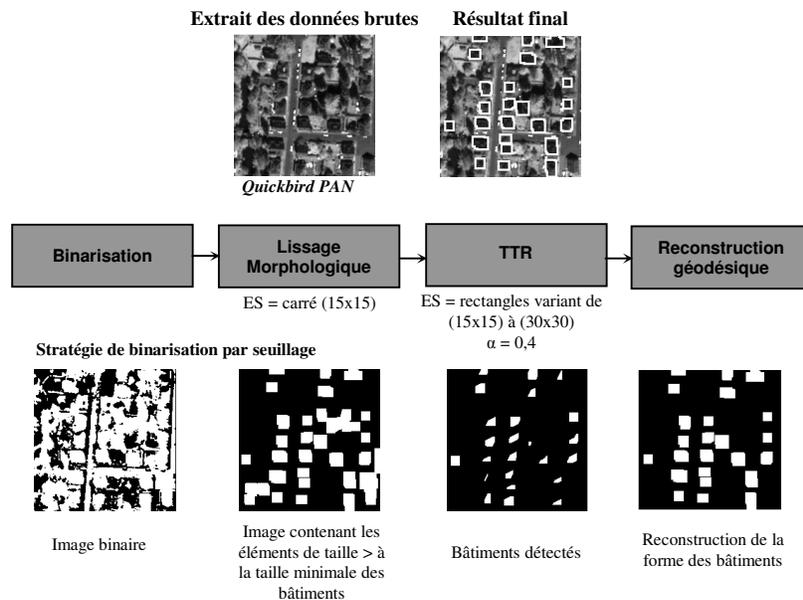


Figure 6. Application de la méthode sur un extrait des données de l'image Quickbird en adoptant une binarisation par seuillage.

Plusieurs remarques peuvent être émises au sujet des résultats illustrés en figure 6. On remarque d'abord que la méthode est effectivement capable de détecter des bâtiments de tailles variables, suite à l'application durant la TTR d'ES rectangulaires de longueur et largeur croissantes. Ensuite, il n'est pas nécessaire que les bâtiments

soient de forme strictement carrée ou rectangulaire pour être détectés. Ils doivent seulement s'en approcher. Les bâtiments présentant de petits décrochements ou des parties cachées sont donc aussi sélectionnés. Ceci est rendu possible grâce à l'introduction de la zone d'incertitude dans la TTR (défini ici avec un coefficient α de 0,4). Cette zone d'incertitude est particulièrement importante puisqu'elle garantit une certaine tolérance sur la forme des objets analysés dans l'image. Une illustration de celle-ci est fournie en figure 7 avec un bâtiment présentant une forme légèrement édentée suite aux différences de réflectance et aux ombres.



Figure 7. Zone d'incertitude introduite dans la TTR en (c) pour détecter un bâtiment n'étant pas parfaitement rectangulaire (image originale en (a) et image binarisée en (b)). La position possible de l'ES de l'objet apparaît en gris clair tandis que celle de l'ES' du fond apparaît en gris foncé. Entre ces deux zones, ce n'est ni purement du bâti (pixels blancs), ni purement du fond (pixels noirs).

Il faut enfin noter que la TTR permet d'éliminer des éléments dans l'image binaire qui ne correspondent pas à des bâtiments. C'est le cas de certaines zones d'ombres notamment. Le nombre d'objets présents dans l'image de la TTR (étape 3) est donc inférieur à celui dans l'image lissée (étape 2). La forme des objets reconstruite lors de la dernière étape correspond quant à elle, à celle des éléments de l'image lissée (image masque).

La méthode a également été appliquée en adoptant l'approche de binarisation par classification non supervisée d'histogramme. Nous illustrons les résultats obtenus en figure 8 sur un autre extrait de l'image, en présentant différents clusters calculés et leur fusion (ici uniquement les couples). Pour rappel, chaque cluster représente un maximum local de l'histogramme de l'image (cf. 3.1.2). A travers ces images, on se rend compte de l'intérêt de réaliser la fusion des clusters. On constate en effet qu'avant celle-ci, aucun cluster ne représentait à lui seul les bâtiments. En fusionnant les clusters, une des images obtenues (le couple 1-3) contient cette fois la presque totalité des éléments constitutifs des bâtiments. La TTR a ensuite été appliquée sur chacun des clusters fusionnés pour obtenir, après union des résultats intermédiaires, le résultat final illustré.

Afin d'évaluer de manière quantitative la qualité de la méthode de détection des bâtiments, des matrices de confusion ont été calculées en prenant en compte l'ensemble des extraits traités (tableau 1). Les données de référence nécessaires au calcul ont été obtenues manuellement, par digitalisation. L'évaluation a été faite par régions, en comptant le nombre de pixels pour chaque région détectée (que ce soit une détection correcte, un faux positif ou un faux négatif).

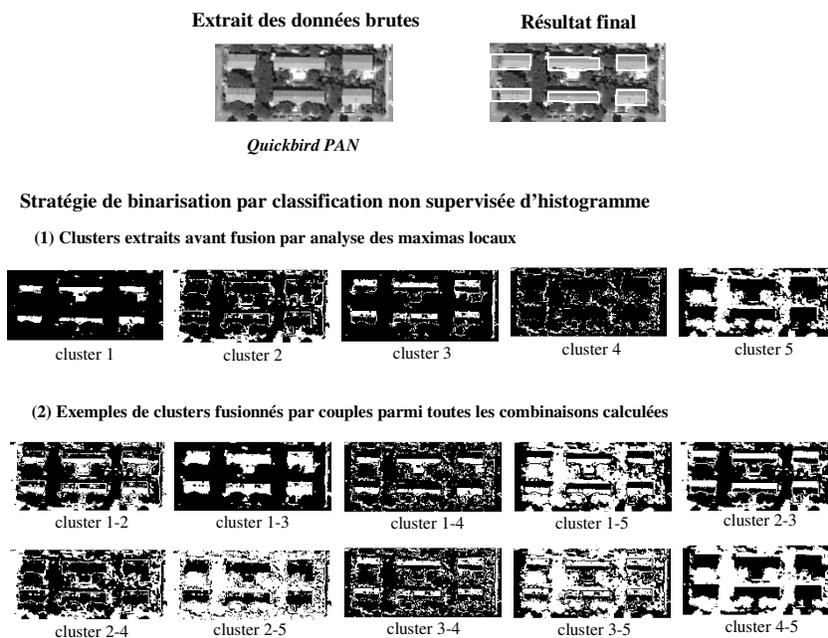


Figure 8. Illustration de la méthode de binarisation par classification non supervisée d'histogramme

Tableau 1. Matrices de confusion calculées en tenant compte de l'ensemble des extraits traités (évaluation par pixels).

	Binarisation par seuillage		Binarisation par classification		
	Bâti	Non bâti	Bâti	Non bâti	
Bâti	18788	1919	Bâti	17673	
Non bâti	9023	92670	Non bâti	10138	
Précision globale : 0,91			Précision globale : 0,88		
Kappa : 0,71			Kappa : 0,63		

Dans le cas d'une binarisation par seuillage, on obtient un taux de précision globale de 0,91 contre 0,89 pour une binarisation par classification. Un certain

nombre de bâtiments dans l'image n'ont donc pas été détectés par la méthode, soit parce qu'ils ont été éliminés à tort lors de l'étape de lissage ou durant la TTR (faux négatifs). Inversement, certains éléments ont été considérés comme des bâtiments alors qu'ils n'auraient pas dû l'être (faux positifs). Bien que la méthode par seuillage fournisse globalement de meilleurs résultats quantitatifs, elle est plus délicate à paramétrer et son usage se limite aux images où les bâtiments sont soit plus clairs, soit plus sombres que leur environnement. Au contraire, l'approche par classification permet de s'affranchir de cette limite, même si dans les cas les plus simples elle ne fournit pas de meilleurs résultats que la première approche.

Les erreurs qui apparaissent dans la détection des bâtiments ont plusieurs origines. Il existe un certain nombre de cas pour lesquels plusieurs bâtiments voisins ont été agrégés durant l'étape du lissage morphologique, avant d'appliquer la TTR. Ces agrégations ont généralement pour effet de produire un élément de grande taille, ayant une forme relativement complexe, qui s'écarte trop de la forme rectangulaire. Dans ce cas, l'élément est éliminé durant la TTR car il ne répond plus aux conditions pour lesquelles sont définis les ES (figure 9). Ces agrégats devraient être traités avec d'autres ES adaptés. L'agrégation ne se fait d'ailleurs pas toujours uniquement entre des bâtiments. Nous avons rencontré des cas pour lesquelles il y avait une agrégation avec des arbres par exemple. Ces arbres ayant des réponses spectrales assez proches des bâtiments dans l'image panchromatique, ceux-ci ne sont pas toujours éliminés durant la binarisation, ce qui explique leur présence dans le masque obtenu avant l'application de la TTR.

Pour améliorer les résultats, il faudrait envisager d'appliquer d'autres ES afin de gérer les agrégats créés. Par ailleurs, certaines erreurs pourraient être sans doute évitées en exploitant des images multispectrales. Elles permettraient par exemple de calculer l'indice de végétation qui pourrait être utilisé pour éviter des confusions avec les arbres.

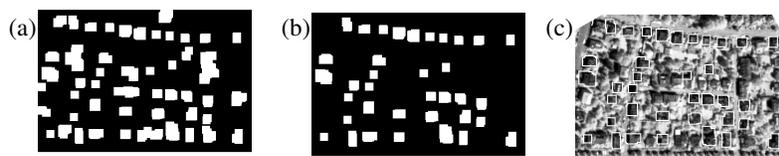


Figure 9. L'agrégation d'éléments lors de l'étape du lissage morphologique a pour effet de générer des objets de forme trop complexe pour pouvoir être gardés durant la TTR. L'image (a) représente le masque obtenu après le lissage. Les images (b) et (c) correspondent au résultat final.

Les matrices de confusion ne donnent aucune indication sur la qualité de la forme des bâtiments extraits. Bien que ceux-ci présentent quelques défauts, nous pouvons

considérer, après un examen visuel, que cette qualité est satisfaisante, comme en témoigne les extraits affichés.

À titre de comparaison, nous donnons en figure 10 le résultat de segmentations obtenues en appliquant deux autres méthodes, à savoir : une ligne de partage des eaux [Vincent et Soille 1991] et une segmentation par croissance de régions [Batz et Schäpe 2000]. La première a été réalisée à partir du calcul d'un gradient morphologique avec un ES carré de 3x3, lissé ensuite par une érosion avec un ES carré de 5x5. La ligne de partage des eaux a été appliquée sur cette image lissée. La seconde méthode de segmentation a été réalisée à l'aide du logiciel eCognition (Definiens-Imaging). Les paramètres optimaux, définis empiriquement par tâtonnements, correspondent à un facteur d'échelle de 25 et un critère d'homogénéité de 0,2 pour la couleur, et 0,8 pour la forme. Ces méthodes de segmentation globale fournissent des régions de formes moins représentatives des bâtiments.

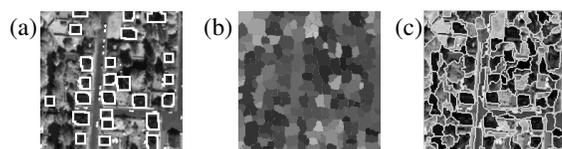


Figure 10. Comparaison des résultats de la méthode proposée en (a) avec ceux de deux segmentations globales différentes : une segmentation par ligne de partage des eaux (b) et une segmentation par croissance de régions (c).

5. Conclusion et perspectives

Nous avons proposé dans cet article, une méthode d'extraction automatique des bâtiments destinée à des images satellitaires THRS. Cette méthode est fondée sur l'utilisation d'opérateurs issus de la morphologie mathématique binaire et en particulier, une adaptation de la transformée en « tout ou rien ». La qualité des résultats obtenus nous permettent de considérer l'approche comme globalement efficace dans le contexte choisi, à savoir : des quartiers pavillonnaires dans lesquels les bâtiments de tailles variables ont des formes proches de rectangles et sont individualisés. L'approche proposée ne se limite cependant pas à ce contexte. La généralité de la méthode devrait permettre d'extraire des bâtiments correspondant à d'autres motifs tels que des bâtiments en « U », « T » ou « L » représentés dans différentes directions. Il suffit pour cela de remplacer l'ES rectangulaire testé par un ES correspondant à la forme recherchée en s'assurant que la méthode reste insensible au bruit.

Concernant les améliorations à apporter, un aspect important de l'approche devrait être revu pour la rendre plus opérationnelle. Il concerne le choix de la taille des ES utilisés lors des étapes du lissage morphologique et de la TTR. Ce choix est

actuellement laissé à l'opérateur qui se charge de repérer les bâtiments les plus petits et les plus grands dans l'image et de calculer leur taille pour définir les ES. Nous envisageons à l'avenir de déterminer ces ES automatiquement, en calculant sur l'image traitée, son profil morphologique par analyse granulométrique [Benediktsson et al. 2003]. Le profil morphologique correspond à l'histogramme des tailles des objets présents dans l'image. Il est obtenu en appliquant successivement des ouvertures avec un ES de taille croissante. Ce profil devrait permettre de détecter les bâtiments les plus petits et les plus grands automatiquement. C'est ce que nous confirment les premières expérimentations déjà réalisées.

Du point de vue de la validation de l'approche, il est également envisageable d'évaluer plus finement et de manière quantitative, la qualité morphologique des bâtiments. Des indices de forme (compacité, élongation...) peuvent en effet être calculés et comparés aux données de référence. L'approche proposée sera enrichie dans ce sens.

Enfin, il serait également intéressant par la suite d'étendre l'approche à la morphologie mathématique fonctionnelle de manière à exploiter des images en niveaux de gris ou des images multispectrales.

6. Remerciements

Cette recherche a bénéficié du soutien financier de l'ACI "Masse de données" (2004-2007) dans le cadre du projet FoDoMuSt (Fouille de Données Multi-Stratégie pour extraire et qualifier la végétation urbaine).

7. Bibliographie

- Baatz M. et Schäpe A. 2000. « Multiresolution segmentation: an optimization approach for high quality multi-scale image segmentation », In *Proceedings of the 12th Angewandte Geographische Informationsverarbeitung*, pp. 12-23.
- Benediktsson J., Pesaresi M. and Arnason K. 2003. « Classification and feature extraction for remote sensing images from urban areas based on morphological transformations: Urban remote sensing by satellite », *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, n°9, pp. 1940-1949.
- Benz U., Hofmann P., Willhauck G., Lingenfelder I. and Heynen M. 2004. « Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information », *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 58, pp. 239-258.
- Brunn, A. and Weidner, U. 1997. « Extracting Buildings from Digital Surface Models », In *Proceedings of the ISPRS Commission III/IV Workshop on 3D Reconstruction and Modelling of Topographic Objects*, Stuttgart, Germany, IAPRS, pp. 27-34.

- Callot Y, Mering C. et Simonin A. 1993. « Quantification et cartographie par télédétection des aires dunaires sur les images SPOT panchromatiques à partir des méthodes de la morphologie mathématique », *Bull. Inst. Géol. Bassin d'Aquitaine*, Bordeaux, vol. 53, pp. 221-232.
- Cocquerez J.-P. et Philipp S. 1995. *Analyse d'images : filtrage et segmentation*. Paris : Masson, 488 p.
- Destival I. 1986. « Mathematical morphology applied to remote sensing », *Acta Astronautica*, vol. 13, n°6-7, pp. 371-385.
- Donnay J.-P., Barnsley M. & Longley P. (eds) 2001. *Remote Sensing and urban Analysis*. Londres: Taylor & Francis.
- Flouzat G. 1988. « Morphologie mathématique et analyse des images de télédétection », *PhotoInterprétation*, vol. 27, n°5, pp. 1-16.
- Jin X. and Davis C. 2005. « Automated Building Extraction from High-Resolution Satellite Imagery in Urban Areas Using Structural, Contextual, and Spectral Information », *EURASIP Journal on Applied Signal Processing*, vol. 14, pp. 2196-2206.
- Jishuang Q. and Chao W. 2002. « A multi-threshold based morphological approach for extracting coastal line feature in remote sensing », In *Proceedings of the ISPRS Commission I Symposium on Integrated Remote Sensing at the Global, Regional and Local Scale*, IAPRS, vol. XXXIV, part 1.
- Katartzis A., Pizurica V. and Salí H. « Application of Mathematical Morphology and Markov random field theory to the automatic extraction of linear features in airborne images », In *Proceedings of the International Symposium on Mathematical Morphology and its Applications to Image and Signal Processing V (ISMM'00)*, Palo Alto, USA, pp. 405-414.
- Laporterie F., Flouzat G. and Amram O. 2001. « Mathematical morphology multi-level analysis of trees patterns in Savannas », In *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS'01)*, Sydney, Australia, pp. 1496-1498.
- Matti-Gallice C. et Collet C 2004. « Morphologie mathématique et échelle : extraction du bâti à différentes résolutions spatiales », *Revue internationale de Géomatique*, vol. 14, n°3-4, pp. 441-463.
- Neubert M., Herold H. and Meinel G. 2006. « Evaluation of remote sensing image segmentation: further results and concepts », In *Proceedings of the 1st International Conference on Object-Based Image Analysis (OBIA'06)*.
- Otsu N. 1979. « A Threshold Selection Method from Grey-Level Histograms », *IEEE Transactions on Systems, Man and Cybernetics*, vol. 9, n°1, pp. 62-66.
- Ridler T., et Calvard S. 1978. « Picture Thresholding Using an Iterative Selection Method », *IEEE Transactions on Systems, Man and Cybernetics*, vol. 8, n°8, pp. 629-632.
- Serra J. 1982. *Image Analysis and Mathematical Morphology*. London : Academic Press, 610 p.

- Soille et Pesaresi 2002. « Advances in mathematical morphology applied to geoscience and remote sensing », *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, n°9, pp. 2042-2055.
- Soille P. 2004. *Morphological Image Processing: Principles and Applications*. Heidelberg: Springer-Verlag, 391 p.
- Vincent L. et Soille P. 1991. « Watersheds in Digital Spaces: An Efficient Algorithm Based on Immersion Simulations », *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, n°6, pp. 583-598.
- Weber C. 1995. *Images satellitaires et milieu urbain*. Paris : Hermès-Lavoisier, 185 p.
- Weber J. Lefèvre S. et Sheeren D. Détection des bâtiments dans les images THRS avec la morphologie mathématique, *Actes du colloque internationale de Géomatique et d'Analyse Spatiale (SAGEO'06)*, Strasbourg, France, résumé long + poster.
- Zhang C., Murai S. and Baltsavias E. 1999. « Road Network Detection by Mathematical Morphology », In *Proceedings of the ISPRS Workshop on 3D Geospatial Data Production: Meeting Application Requirements*, Paris, France, pp. 185-200.
- Zheng X., Gong P. and Strome M. 1995. « Characterizing spatial structure of tree canopy using color photographs and mathematical morphology », *Canadian Journal of Remote Sensing*, vol. 21, n°4, pp. 420–428.