CLASSIFICATION ORIENTÉE OBJET D'UNE IMAGE *QUICKBIRD* DE LA VILLE DE HANOI (VIÊTNAM) : ÉTUDE EMPIRIQUE DES PARAMÈTRES DE SEGMENTATION

Thi-Thanh-Hiên Pham

[a] *Département d'études urbaines et touristiques, Université du Québec à Montréal,*
315, rue Sainte-Catherine Est, Montréal, Québec, H2X 3X2
Courriel : pham.thi_thanh_hien@uqam.ca
Téléphone : (514) 987-3000 poste 5540

Dong-Chen He

Département de géomatique appliquée et CARTEL, Université de Sherbrooke,
2 500, boulevard de l'Université, Sherbrooke, Québec, Canada J1K 2R1
Courriel : Dong-Chen.He@USherbrooke.ca
Téléphone : (890) 821 8000 poste 62299

Soumis le 25 septembre 2009 ; accepté le 30 septembre 2012 - © *Revue Télédétection*, 2012, vol. 11, n° 2, p. 287-305

Résumé

L'arrivée de l'imagerie satellite à très haute résolution spatiale (THRS) ouvre de nouvelles applications en télédétection urbaine. Malgré les nombreux avantages de l'approche orientée objets pour le traitement de l'image à THRS en milieu urbain, les paramètres de réalisation de cette approche sont encore peu documentés. L'objectif de l'article est de déterminer de façon empirique les paramètres de segmentation pour l'extraction des objets urbains à partir d'une image Quickbird (0,7 m). L'étude porte sur un milieu urbain dense, soit un des districts centraux de la ville de Hanoi (Viêtnam). Les tests empiriques sont effectués sur les paramètres de segmentation dans eCognition, notamment les ratios de couleur/forme et de compacité/lissage. Pour les classes d'eau et de végétation (zones agricoles, arbres de parcs, arbres au bord des rues, arbres isolés), la couleur est nettement plus importante que la forme et ce, avec une faible compacité de segmentation. L'extraction d'une grande route est satisfaisante avec une faible compacité, mais les petites rues ne sont pas adéquatement classifiées. Pour les îlots de bâtiments, la forme est plus pertinente que la couleur et la compacité est plus importante que le lissage. Afin d'améliorer la classification des arbres dispersés et des petites rues, nous recommandons d'utiliser des données d'altitude ou de plus haute résolution spectrale ainsi que des méthodes de traitement d'images plus adaptées.

Mots clés : *approche orientée objet, Quickbird, paramètres de segmentation, milieu urbain*

Abstract

Very high spatial resolution (VHSR) images have led to the development of new applications in urban remote sensing. Despite the many advantages of the object oriented approach, the exact parameters that should be used in processing VHSR images for urban settings are still ill-documented. In this paper, we determine empirically the best parameters of segmentation for the extraction of urban objects from a Quickbird image (0.7m). The study is conducted in a dense central district of Hanoi (Vietnam). Empirical tests on segmentation parameters are conducted in eCognition, with a specific focus on the ratios of colour/shape and of compactness/smoothness. For classes of water and vegetation (agriculture, street trees, park trees and isolated trees), colour is much more important than shape, and compactness is less important than smoothness. Extraction of a large road is successfully achieved by using a weak compactness, but small streets are not adequately classified. For housing blocks, shape is more important than colour while compactness is more important than smoothness. In order to improve the classification of isolated trees and small streets, we recommend to use altitude data or data with higher spectral resolution, along with methods that are more adapted to those classes.

Keywords: *object-oriented approach, Quickbird, segmentation parameters, urban settings.*

1. INTRODUCTION

L'arrivée de l'imagerie satellite à très haute résolution spatiale (THRS), comme Ikonos, Quickbird ou WorldView, ouvre de nouvelles applications en télédétection. Elle permet de favoriser la mise à jour des bases de données spatiales, ce qui est primordial en milieu urbain (Longley *et al.* 2001). De telles bases de données sont extrêmement importantes dans les pays en développement où l'urbanisation progresse rapidement (United Nations 2008) et où les sources de données conventionnelles sont imprécises, difficiles à mettre à jour, et d'accès limité (Miller and Small 2003). La recherche en télédétection urbaine s'oriente donc vers l'extraction d'objets urbains à partir de nouveaux capteurs d'images à très haute résolution spatiale (par exemple Bhaskaran *et al.* 2010; Herold *et al.* 2002; Herold *et al.* 2003).

Le milieu urbain est extrêmement hétérogène tant au niveau spatial qu'au niveau spectral. La complexité spatiale se manifeste par la densité et la diversité des objets s'y trouvant. Au niveau spectral, la variabilité spectrale à l'intérieur d'une catégorie d'objets (ou d'une classe d'occupation du sol) augmente quand la taille de pixel est réduite, comme dans le cas des capteurs à THRS. Par exemple, il devient possible de distinguer les matériaux avec lesquels les toits sont construits grâce à leurs couleurs. Cependant, les caractéristiques spectrales des capteurs THRS actuels sont insuffisantes pour la cartographie du milieu urbain (Herold *et al.* 2002). Herold *et al.* (2003) ont montré que même avec une quantité importante d'informations spectrale obtenues à partir des 224 bandes d'un capteur hyper-spectral (AVIRIS), la classification par pixels n'est pas capable de séparer les classes d'occupation des sols urbains. Cela confirme que l'information spectrale à elle seule, y compris celle obtenue avec des images THRS, est insuffisante pour discriminer les objets urbain (Guindon 1997). De nouvelles études se sont donc penchées sur une autre approche appelée « orientée objet », qui intègre d'autres types informations que les signaux spectraux et qui est basée sur la création d'objets d'image (Hay and Castilla 2006). Plusieurs auteurs montrent que cette approche donne de meilleures cartographies des objets urbains (Bhaskaran *et al.* 2010; Thomas *et al.* 2003; Yu *et al.* 2006; Yuan and Bauer 2003).

L'approche orientée objet comprend deux étapes principales : la création des objets images, qui est appelée segmentation, et leur classification. La segmentation vise à créer des zones individuellement définies en termes d'homogénéité spectrale, texturale et spatiale (la simplicité géométrique du contour des segments) ou selon des catégories thématiques spécifiques à l'étude. Une fois les objets définis, la classification vise à les regrouper à l'aide d'attributs spectraux (ex : moyenne des niveaux de gris de l'objet), spatiaux (ex : taille de l'objet, critère de forme), structuraux (relation spatiale entre les objets) ou d'autres variables géographiques (ex : altitude, pente, ou densité de population). Le résultat final de l'analyse de l'image dépend grandement de la qualité de la segmentation (Dey *et al.* 2010; Taubenböck *et al.* 2006). La segmentation attire ainsi l'attention de plusieurs chercheurs. Au cours des 40 dernières années, environ 4300 algorithmes de segmentation ont été développés et il existe différentes façons de classer ces algorithmes en catégories (Zhang 2006), qui peuvent être divisés en trois groupes: détection des régions, détection des contours (ou arêtes) et méthodes hybrides (Bonn and Rochon 1992; Cheng *et al.* 2001; Dey *et al.* 2010; Guindon 1997; Zhang 1997).

Le premier groupe, détection des régions, cherche à identifier des objets qui partagent certaine similarité des propriétés de niveaux de gris, de couleur ou de texture. Les méthodes comme la croissance de région, le seuillage multi-niveau ou la décomposition-fusion (*split-merge* en anglais) font partie de ce groupe (par exemple Carleer *et al.* 2005; Wuest and Zhang 2009). Le deuxième groupe, *détection des contours*, est basé sur la discontinuité des propriétés. La morphologie mathématique, notamment la transformation de bassin versant ou l'extraction du squelette appartient à ce groupe de méthodes (par exemple Zhang and Couloigner 2006). Finalement, le troisième groupe, *méthode hybride* (ou *coopérative*) de deux groupes de méthodes précédentes, se basent sur les lignes de contour pour faciliter le classement et utilisent l'intensité des pixels pour affiner les lignes de contour des objets à séparer (par exemple Sebari and He 2009).

eCognition, un logiciel commercial pour l'approche orientée objet a été lancé en 2003. La segmentation dans eCognition est un type de détection des régions, plus spécifiquement la propagation ou croissance de région. La formation des régions y est réalisée selon des critères prédéfinis selon l'analyste en employant les informations de couleur, de texture ou de forme. La segmentation exige quatre paramètres que l'utilisateur doit définir : l'échelle de segmentation, le poids de la bande spectrale, les ratios de couleur/forme et de compacité/lissage des objets (Benz *et al.* 2004). Comme les algorithmes ainsi que l'effet des interactions des paramètres ne sont pas documentés dans le guide du logiciel, les auteurs des travaux antérieurs déplorent souvent le manque de détail sur les paramètres utilisés par les autres chercheurs pour leur segmentation (Caprioli and Tarantino 2003; Flanders *et al.* 2003; Sugumaran and Voss 2007), mais tendent à donner eux-mêmes un minimum de paramètres. Le manque de transparence des algorithmes utilisés rend nécessaire des tests et des études plus poussées pour trouver les paramètres optimums de segmentation. Cet inconvénient rend le processus de segmentation par eCognition intuitif et moins transférable d'une étude à l'autre.

Bien que le logiciel eCognition soit accessible et facile à utiliser pour des utilisateurs qui ne maîtrisent pas des outils de programmation, il est important d'établir une procédure de segmentation efficace dans ce logiciel. L'objectif principal de cet article est donc de déterminer de façon empirique les paramètres de segmentation en utilisant le logiciel eCognition pour l'extraction des classes d'occupation du sol d'un milieu urbain dense, soit la ville de Hanoi (Viêtnam). Nous visons quatre classes : la végétation urbaine (avec différentes sous-classes), les plans d'eau, le réseau routier et les bâtiments résidentiels. Nous effectuons d'abord les tests de segmentation de manière empirique, puis les objets produits

par les segmentations testées sont classifiés par la fonction d'appartenance du logiciel. Deux validations sont ensuite effectuées pour évaluer l'exactitude de la méthode. Finalement nous discutons quelques limitations de cette données pour une ville aussi dense que Hanoi et suggérer des pistes de recherche pour d'autres types de milieux urbains.

2. MÉTHODOLOGIE

2.1. Données

Nous avons utilisé une image Quickbird de la ville de Hanoi (Viêt-Nam) acquise le 5 novembre 2002. L'image est constituée de quatre bandes spectrales (2,8 m de résolution spatiale): bleu (0,45-0,52 µm), vert (0,52-0,60 µm), rouge (0,63-0,69 µm) et proche infrarouge (0,76-0,90 µm)) fusionnées avec la bande panchromatique (0,7m de résolution spatiale).

2.2. La zone d'étude et les classes d'occupation du sol

Nous avons choisi Hoàn Kiếm, un des districts centraux de la ville de Hanoi (Viêtnam), comme zone d'étude. Le milieu urbain vietnamien connait de nombreux problèmes typiques des pays en développement, soient une urbanisation rapide, une concentration démographique importante (Dang and Le 2004; Drakakis-Smith and Kilgour 2001), des problèmes environnementaux et un retard du système de gestion urbaine (O'Rourke 2005). Dans ce contexte, une base de données géographique pour une gestion efficace du milieu urbain au Viêtnam devient cruciale.

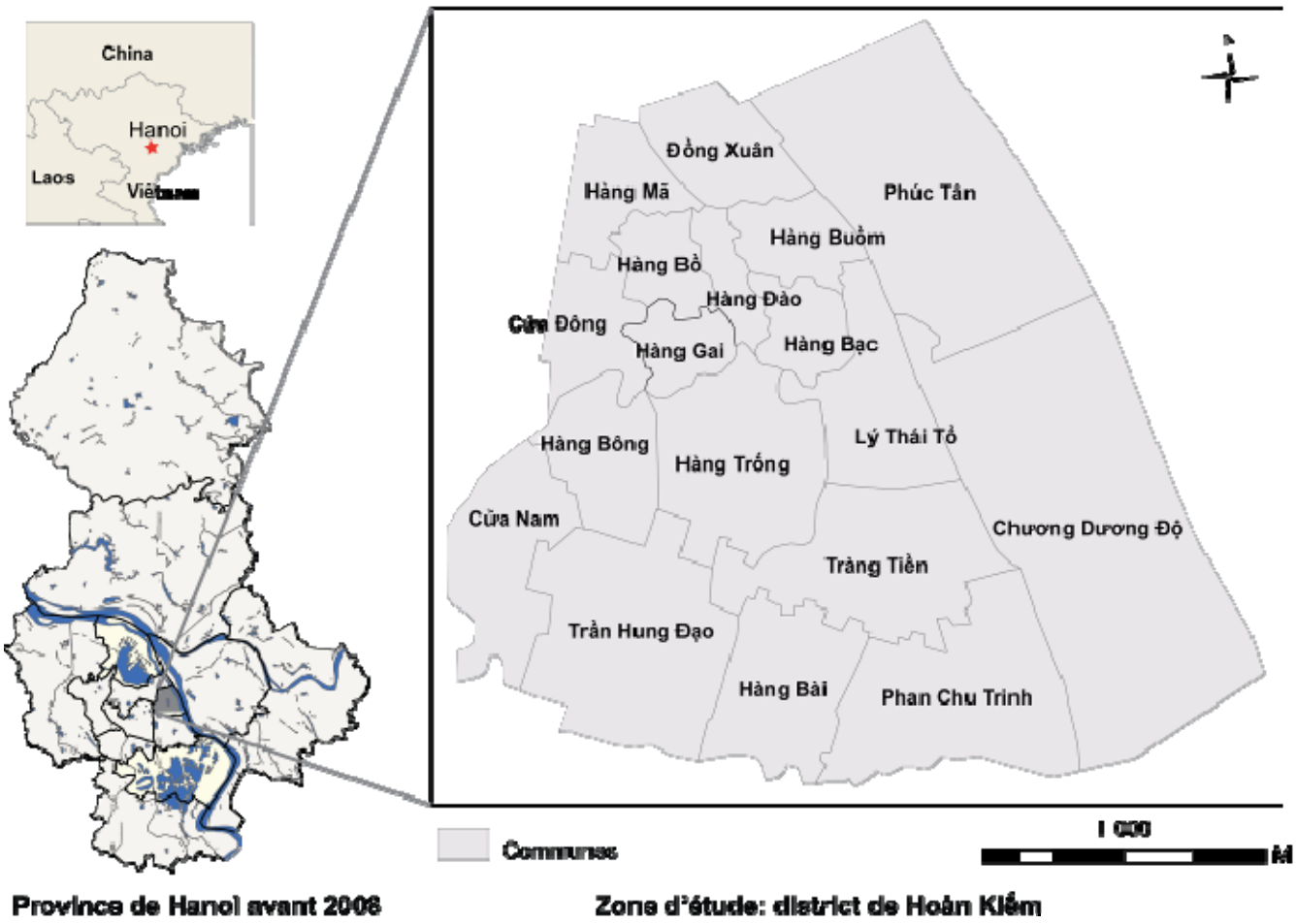


Figure 1. Localisation de la province de Hanoi et la zone d'étude
Figure 1. Location of the province of Hanoi and the study area

Le district Hoàn Kiếm, occupant une superficie de 5,3 km^2 et comprenant 18 communes, se situe au centre historique de la capitale (Figure 1). En termes de caractéristiques architecturales, la zone d'étude comprend trois quartiers. Le premier est le plus ancien de la ville, caractérisé par des maisons tubes appelées compartiments chinois datant majoritairement du XIX^e siècle. Le quartier, très dense en terme de concentration démographique, manque critiquement d'espaces ouverts et de verdure (Procacci and Luong 2007; Tô 2008). Le deuxième a été construit à l'époque coloniale (1888-

1954) sur les terres concédées à la France sous le protectorat (avant 1888). Il se trouve dans le sud du district et est caractérisé par un plan en damier, avec des rues régulières et plantées d'arbres (Procacci and Luong 2007). Le dernier quartier, situé à l'est du district, est un quartier récent qui a été établi dans les années 1990 sur des terres jusque là occupées par des logements spontanés illégaux. Des terres agricoles se trouvent à l'est du quartier, au bord du fleuve Rouge.

Nous visons à obtenir les classes d'occupation du sol représentatives de la zone d'étude: plans d'eau, agriculture, végétation urbaine (comprenant les arbres dans les parcs, les arbres au bord des rues et les arbres individuels), routes et bâtiments résidentiels. Il est à noter que comme la ville est très dense, le sol nu y existe mais reste rare. Nous ne l'avons considéré dans la classification. Avant d'effectuer le processus de traitement de l'image, nous avons identifié les caractéristiques de couleur et de forme de ces classes. Les plans d'eau (incluant des lacs et la fleuve Rouge) sont facilement distingués sur l'image grâce à leur valeur spectrale basse dans toutes les bandes et à leur taille relativement grande. Différents types de végétation se trouvent dans la zone d'étude: des zones agricoles, des groupes d'arbres de forme polygonale (les parcs municipaux), des groupes d'arbres ayant une forme allongée (qui se trouvent souvent au bord des rues ou sur les trottoirs) et des arbres individuels. Les rues, quant à elles, sont principalement de couleur grise mais avec plusieurs nuances variant de pâle à foncé à cause des différent matériaux utilisés pour recouvrir la chaussée. Les objets se trouvant au bord ou sur les rues, comme les arbres, les toits des maisons et les véhicules ont cependant des réponses spectrales différentes. Ces variations de couleur rendent l'extraction des rues difficile, voire parfois impossible. Les îlots résidentiels sont composés de maisons souvent de forme carrée ou rectangulaire, de petite taille (environ de 10m de chaque côté) et entassées les unes sur les autres. Elles sont souvent entourées par des rues ou des arbres. Les descriptions de ces classes serviront à déterminer les valeurs de paramètres de segmentations.

2.3. Segmentation

2.3.1. Segmentation par eCognition

La segmentation par eCognition commence par le niveau spatial le plus bas : les objets de départ sont les pixels se trouvant sur l'image. Les pixels sont ensuite regroupés par plusieurs paramètres qui garantissent que l'hétérogénéité spectrale et spatiale du groupe est minimale (Benz *et al.* 2004). Lors de la fusion de deux objets en un groupe, l'hétérogénéité f du nouvel objet (appelé groupe dans les équations suivantes) est calculée comme suit :

$$f = w_{couleur}.\Delta h_{couleur} + w_{forme}.\Delta h_{forme} \quad [1]$$

Où wcouleur et wforme sont le poids des propriétés spectrales et de la forme. Ils sont définis par l'utilisateur. Leur somme est égale à 1. Le poids de couleur est privilégié quand l'objet à extraire a une réponse spectrale distincte, comme pour les plans d'eau et les groupes d'arbres (Flanders *et al.* 2003; Laliberte *et al.* 2004; Lang *et al.* 2006).

$\Delta h_{couleur}$ *est calculé pour une bande c ayant le poids wc, comme dans l'équation suivante :*

$$\Delta h_{couleur} = \sum_c w_c \left(n_{groupe}.\sigma_{c,groupe} - (n_{obj1}.\sigma_{c,obj1} + n_{obj2}.\sigma_{c,obj2})\right) \quad [2]$$

Où n est le nombre de pixels contenus dans un groupe ou un objet et σ est l'écart-type de la valeur spectrale c du groupe et des objets. Le poids wc est défini par l'utilisateur. eCognition accepte des bandes matricielles et vectorielles.

Δh_{forme} *est calculé :*

$$\Delta h_{forme} = w_{compt}.\Delta h_{compt} + w_{liss}.\Delta h_{liss} \quad [3]$$

Où les poids wcompt et wliss sont définis par l'utilisateur. Le poids de compacité est privilégié quand l'objet à extraire a une forme compacte, par exemple un rectangle (Flanders *et al.* 2003). Δh_{compt} est l'hétérogénéité de compacité qui est calculée par la longueur l du contour de l'objet :

$$\Delta h_{compt} = n_{groupe}.\frac{l_{groupe}}{\sqrt{n_{groupe}}} - (n_{obj1}\frac{l_{obj1}}{\sqrt{n_{obj1}}} + n_{obj2}\frac{l_{obj2}}{\sqrt{n_{obj2}}}) \quad [4]$$

Δh_{liss} est l'hétérogénéité de lissage qui est calculée par le ratio de la longueur l du contour de l'objet sur la longueur b du rectangle entourant l'objet, comme suit :

$$\Delta h_{liss} = n_{groupe} \cdot \frac{l_{groupe}}{\sqrt{b_{groupe}}} - (n_{obj1} \frac{l_{obj1}}{\sqrt{b_{obj1}}} + n_{obj2} \frac{l_{obj2}}{\sqrt{b_{obj2}}}) \quad [5]$$

Les algorithmes de segmentation d'eCognition minimisent l'hétérogénéité calculée par les équations ci-dessus jusqu'à ce que le nouvel objet ait la taille désirée. La taille désirée est définie par l'utilisateur grâce au paramètre d'échelle.

Il est possible aussi d'effectuer des segmentations hiérarchiques (multi-échelle) dans eCognition. En analyse d'image numérique, la hiérarchie image est une combinaison d'analyses à différentes échelles (par exemple, différentes échelles de segmentation de l'image). Elle permet de se concentrer seulement sur les zones d'intérêt en commençant l'analyse avec une segmentation à résolution grossière, puis en utilisant une segmentation plus fine sur les zones de l'image correspondant à des critères prédéfinis (Kropatsch *et al.*, 2001). Dans eCognition, la hiérarchie image peut être utilisée dans deux étapes du processus de traitement, soit la segmentation multi-échelle et l'établissement de classes. Lors de l'établissement des classes, il faut définir les classes elles-mêmes ainsi que les relations verticales et horizontales entre les classes. Pour simplifier le travail, la hiérarchie des classes peut permettre aux classes filles d'hériter des propriétés de leur classe mère.

2.3.2. Segmentations effectuées dans cette étude

Dans cette étude, nous avons tout d'abord testé les combinaisons de bandes de l'image et identifié celles qui résultent en une meilleure séparabilité des classes qui nous intéressent. La séparabilité des classes a été basée d'abord sur une évaluation visuelle et ensuite empiriquement sur des attributs spectraux et texturaux des objets segmentés (les attributs sont listés dans le Tableau 1). Il est noté que nous avons aussi ajouté la bande NDVI (*Normalized Difference Vegetation Index*) en plus des bandes spectrales d'origine de l'image. L'indice NDVI a été calculé par l'équation suivante :

$$NDVI = \frac{PIR - R}{PIR + R} \quad [6]$$

Où PIR et R sont les bandes de proche infrarouge et rouge de l'image Quickbird.

La bande de NDVI a été utilisée pour la segmentation des classes de végétation et d'eau. Nous avons ajouté une bande raster contenant les classes de végétation et une bande vectorielle contenant le réseau routier pour la segmentation de la classe « îlots de bâtiments ».

Nous avons ensuite choisi l'échelle à préciser appropriée pour chacune des classes. Nous avons commencé la segmentation de l'image avec des échelles grossières (par exemple : 500, 200, 100) et nous sommes ensuite passés à des échelles plus fines (de 50 à 5). Notez qu'au début des tests, le ratio de couleur/forme a été provisoirement fixé à 1:0 (en privilégiant l'homogénéité de couleur des objets) et le ratio compacité/lissage à 0:1 (en privilégiant le lissage de contour des objets). Ces ratios ont été ensuite graduellement variés suivant la méthode décrite à la fin de cette section.

Pour les classes de végétation, nous avons utilisé une démarche de segmentation multi-échelle. Dans un premier temps, une segmentation à l'échelle 100 nous a servi à classifier trois classes : « parc », « autres 1 » et « agriculture 1 » qui correspond aux champs de légumes se trouvant sur la dune de sable sur le fleuve. Nous avons ensuite classifié les objets appartenant à « autres 1 » en trois nouvelles classes : « arbres de rue », « autres 2 » et « agriculture 2 » qui correspond aux champs de légumes se situant au bord du fleuve, dont la taille est plus petite que ceux assignés dans la classe « agriculture 1 ». Finalement, une segmentation à l'échelle 25 sur les objets de classe « autres 2 » nous a permis de distinguer la classe « arbres isolés » du reste des objets. Pour la classe « route », une séquence de segmentation particulière a été utilisée. Une segmentation à échelle fine (à 15) a été d'abord effectuée et ensuite une ré-segmentation à 50 à partir de la segmentation à 15 fusionnée.

Une fois obtenue une combinaison de bandes et une échelle acceptable pour chaque classe, les meilleures valeurs pour les ratios de couleur/forme et de compacité/lissage ont été ajustées. Pour obtenir les meilleures valeurs, nous avons essayé toutes les possibilités en faisant varier la valeur d'un des ratios tout en conservant l'autre constant. Pour chacun des deux ratios, les valeurs des numérateur et dénominateur peuvent varier de 0 à 1, mais leur somme doit toujours être égale à 1. Nous avons donc testé toutes les valeurs en les faisant augmenter par intervalles de 0,2 (0/1; 0,2/0,8; 0,4/0,6; 0,6/0,4; 0,8/0,2; 1/0). Pour chacun de ces ratios, nous avons appliqué une classification avec des règles simples (à mentionner dans la section 2.4.). Chaque classification a été superposée à l'interprétation visuelle numérisée afin d'estimer l'exactitude de la segmentation (voir la section 2.5.). Les segmentations ayant la meilleure exactitude ont été retenues et utilisées pour une classification plus élaborée.

2.4. Classification

La classification par eCognition est réalisée par l'approche du plus proche voisin ou par une fonction d'appartenance définie par l'utilisateur. Dans les premières années suivant le lancement du logiciel, l'approche du plus proche voisin était utilisée dans la plupart des travaux de recherche en classification orientée objet. Cependant, il n'est recommandé d'utiliser cette approche que quand la classe peut être distinguée par un ou peu d'attributs (Baatz *et al.* 2004) et son

exactitude est nettement inférieure à celle produite par l'approche d'appartenance (Platt and Rapoza 2008). Il y a trois groupes d'attributs d'objets que nous pouvons utiliser (Baatz *et al.* 2004) : de réponse spectrale, texturaux, géométriques et contextuels (Tableau 1). La classification utilisant un grand nombre d'attributs est le point le plus avantageux de l'approche orientée objet.

Tableau 1. Attributs utilisés lors de la classification des objets
Table 1. Features used in classification rules

Attributs spectraux	**Attributs texturaux**	**Attributs géométriques**	**Attributs contextuels**
Moyenne (valeur spectrale moyenne d'une bande) Brillance (somme des valeurs spectrales de toutes les bandes) Écarte-type Contraste NDVI Coloration Teinte	Homogénéité Dissimilarité Second moment angulaire Entropie	Superficie Largeur Longueur/largeur Périmètre Courbure Orientation	Distance à la classe « x » Existence de la classe « x » Contour relatif partagé avec la classe « x »

Ecart-type et contraste sont considérés comme des attributs spectraux dans le logiciel eCognition.

La moyenne C_L est la somme des valeurs C_{Li} de tous les pixels composant l'objet, divisée par le nombre de pixels (n).

$$\overline{C_L} = \frac{1}{n}\sum_{i}^{n} C_{Li} \qquad [7]$$

La brillance est calculée par la somme des moyennes $\overline{C_i}$ de toutes les bandes divisée par le nombre de bandes n_L.

$$b = \frac{1}{n_L}\sum_{i}^{nL} \overline{C_i} \qquad [8]$$

L'écart-type est la somme des écart-type de tous les pixels composant l'objet.

$$\sigma_L = \sqrt{\frac{1}{n-1}\sum_{i}^{n}(C_{Li} - \overline{C_L})^2} \qquad [9]$$

L'indice de contraste entre voisins correspond à la valeur différentielle par rapport à un attribut de deux objets adjacents.

L'indice de coloration est souvent utilisé pour classifier les types de sols à partir d'une image satellite car il caractérise des couleurs de sols (Coulibaly and Gwyn 2005). Il est calculé comme suit :

$$Coloration = \frac{(R-B)}{R} \qquad [10]$$

Où R et B sont respectivement la réponse spectrale des bandes rouge et bleue.

L'indice de teinte caractérise les nuances de couleur (Coulibaly and Gwyn 2005). Il est calculé comme suit :

$$Teinte = \frac{\tan^{-1}((R.2) - V - B))}{3.0,5.(V-B)} \qquad [11]$$

Où V, B et R sont respectivement la réponse spectrale des bandes verte, bleue et rouge.

2.5. Évaluations d'exactitude

Nous avons effectué deux validations des résultats de la classification, soit une validation par région dans des zones présélectionnées et une validation par points sur l'image entière. Une validation par région est plus efficace qu'une validation par pixels (Congalton and Green 2009) et la plupart des études sur l'approche objet utilisent ce type de

validation (Carleer *et al.* 2005; Platt and Rapoza 2008; Sebari and He 2009). Par contre, comme ce type de validation est laborieux, notamment à l'étape de l'interprétation visuelle des objets sur l'image, nous avons seulement procédé à la validation par région dans sept zones pré-sélectionnées (Figure 2). Ces zones ont été choisies en basant sur les observations sur le terrain. Dans chaque zone, nous avons essayé de localiser tous les types d'objet appartenant aux classes à identifier. Par exemple, pour la classe « îlots de bâtiment », le but était de classer tous les îlots de bâtiments de type compartiment chinois dans le Vieux quartier, les bâtiments coloniaux dans le quartier français et les bâtiments nouvellement construits dans le nouveau quartier bordant fleuve Rouge. Afin de réduire le temps d'interprétation, chaque zone de test comprend au moins deux classes. La zone 3, par exemple, comprend des îlots de bâtiments (compartiments chinois et bâtiments nouvellement construits) et l'autoroute. Ensuite, nous avons numérisé tous les objets et les avons assignés à une des classes. Nous avons utilisé le logiciel ArcGIS pour comparer cette interprétation visuelle et la classification de eCognition pour obtenir une première matrice de confusion. À partir de cette matrice, nous avons évalué l'exactitude du producteur (*producer accuracy*) et du consommateur (user accuracy), ainsi que l'exactitude moyenne (la moyenne des deux valeurs d'exactitude).

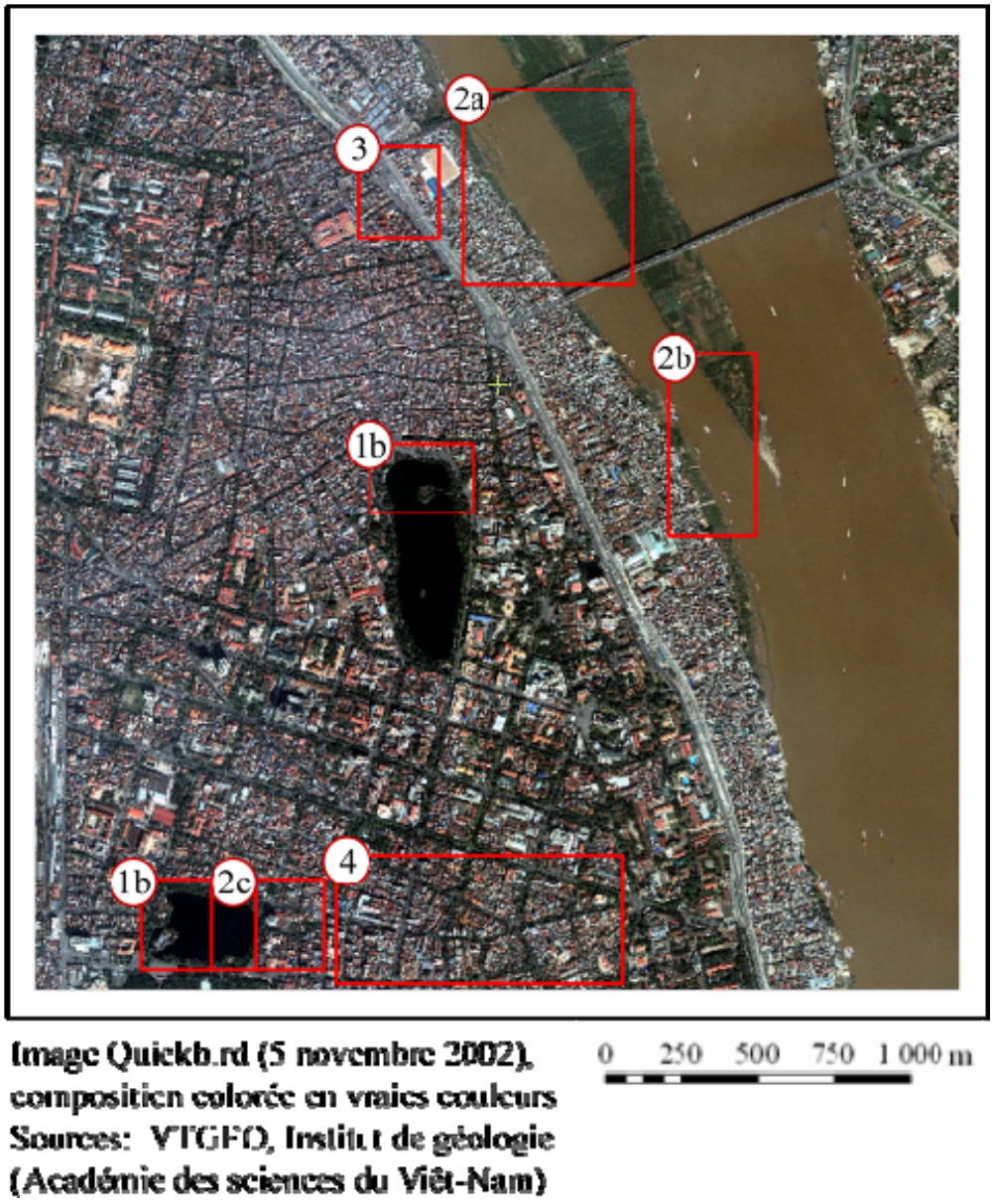


Figure 2. Localisation des zones de validations sur l'image: Zones 1a et b comprenant l'eau de lac et des arbres de parcs; Zones 2a-c comprenant des zone agricoles, des arbres de rues, l'eau du fleuve et des bâtiments; Zone 3 comprenant l'autoroute et des bâtiments; Zone 4 comprenant des bâtiments, des arbres de rues et des arbres isolés.

Figure 2. Location of validation zones: Zones 1a and b including lake water and park trees; Zones 2a-c include agriculture, street trees, river water and residential buildings; Zone 3 including the highway and residential buildings; Zone 4 including residential buildings, street trees and isolated trees

Pour la deuxième validation, nous avons créé un échantillonnage aléatoirement stratifié de 272 points qui sont aléatoirement distribués sur l'image (à l'aide du logiciel ArcGIS). Il est à noter que ces points se trouvent dans tous les classes obtenues par la classification. Comme l'exactitude des classes « eau de lac » et « eau du fleuve » est élevé dans la première validation, elles ont été fusionnées en une seule classe « eau » lors de la deuxième validation. Au total, nous avons créé des points de référence pour les sept classes suivantes : eau, zones agricoles, arbres isolés, arbres de rues, arbres de parcs, route et îlots résidentiels. Nous nous sommes assurés d'avoir au moins 30 points par classe et plus de 30 pour les classes qui occupent une grande superficie de la zone d'étude. Ce type de validation nous permet d'évaluer la qualité de la classification sur toute la zone d'étude. Chaque point a ensuite été attribué à une classe correspondant à notre interprétation de deux pixels l'entourant. Nous avons finalement superposé les points sur la meilleure classification de chaque classe afin de créer une matrice de confusion.

3. RÉSULTATS

Dans cette section, nous présentons les meilleures valeurs obtenues pour chaque paramètre de segmentation ainsi que les résultats des deux validations. La matrice de confusion de la deuxième validation (par point) est présentée dans le tableau 2. L'exactitude globale est de 78%. L'indice de Kappa est de 0,74, ce qui est satisfaisant pour une classification de ce type (Congalton and Green 2009).

Table 2. Matrice de confusion produite par la deuxième validation (unité : point).
Table 2. Confusion matrix produced by the second validation (unit : point)

		Classification									
		Eau	Zones agricoles	Arbres isolés	Arbres de rues	Arbres de parcs	Ilots résid.	Route	Non-classifié	Total	Exact. product. (%)
Référence	Eau	**29**							*2*	31	94
	Zones agricoles		**25**		**1**				*5*	31	81
	Arbres isolés		**1**	**13**	**15**		**2**	**5**	*1*	37	35
	Arbres de rues			**1**	**43**	**1**	**1**	**1**	*4*	51	84
	Arbres de parcs				**7**	**14**			*9*	30	47
	Ilots de résid.						**32**		*2*	34	94
	Route							**55**	*3*	58	95
	Total	29	26	14	66	15	35	61	*26*	272	
	Exactitude consommateur (%)	100	96	93	65	93	91	90			78

Valeurs en gras : nombre de points correctement détectés par la classification et l'interprétation.
Valeurs en italique : nombre de points de référence qui ne sont pas détectés par la méthode de classification

3.1. Extraction de la classe "eau"

La segmentation a été réalisée à une échelle de 100 sur les quatre bandes de l'image Quickbird. Comme la valeur spectrale de l'eau est très basse et donc distinguée sur l'image, seule la valeur spectrale est nécessaire pour distinguer l'eau sur l'image. Le paramètre de couleur a donc été fixé à 1. Quant au ratio de compacité/lissage, nous avons testé trois valeurs de compacité (à 0, 0,5 et 1). L'exactitude des trois tests est légèrement différente.

Selon la première validation, avec la segmentation réalisée par les valeurs de 1/0 et 0/1 pour les ratios de couleur/forme et de compacité /lissage, l'extraction des classes « lacs » et « fleuve » a une exactitude moyenne de 0,97 et 0,98, respectivement. L'extraction de cette classe a donc une bonne précision. Selon la deuxième validation, les deux exactitudes sont aussi satisfaisantes, étant supérieurs à 90%.

3.2. Extraction de la classe « végétation »

Nous avons choisi trois échelles différentes après en avoir évalué plusieurs, soient 100 pour les zones agricoles, 50 pour les arbres de parc, et 25 pour les arbres de rue et les arbres isolés. Nous avons ensuite effectué les tests de combinaison de bandes (Tableau 3) qui montrent que la segmentation sur la bande NDVI donne la meilleure exactitude.

Tableau 3. Exactitude de segmentations avec différentes combinaisons de bandes : classe de « végétation » avec couleur à 1 et compacité à 0 (unité : %)
Table 3. Accuracy of segmentations with different band combinations: « vegetation » class, with coulor = 1 and compacité = 0 (unit : %)

4 bandes Quickbird	**NDVI**	**Exactitude moyenne**	
		Arbres de parc	Arbres de rue
Non	Oui	0,87	0,71
Oui	Oui	0,87	0,67
Oui	Non	0,87	0,67

Afin de connaître l'impact du ratio de compacité/lissage sur la combinaison des bandes, nous avons aussi testé d'autres segmentations sur quatre bandes originales avec une valeur de couleur de 0,3 et des valeurs de compacité variées (Tableau 4). Les niveaux d'exactitude des classes varient peu : 87-88 % pour les arbres de parc et 66-67 % pour les arbres de rue, ce qui suggère que la variation des ratios n'influence pas le choix de bandes. L'exactitude la plus élevée a été obtenue avec une valeur de compacité de 0,5, mais elle est moindre que celle obtenue avec la bande NDVI. Nous avons décidé d'opter pour la bande NDVI parce que le test réalisé avec celle-ci est plus exact et requiert un temps de traitement inférieur. Comme la forte valeur de couleur du ratio de couleur/forme semble propice pour la segmentation de la végétation sur l'image, nous avons fixé ce ratio à 1/0 pour les prochaines segmentations.

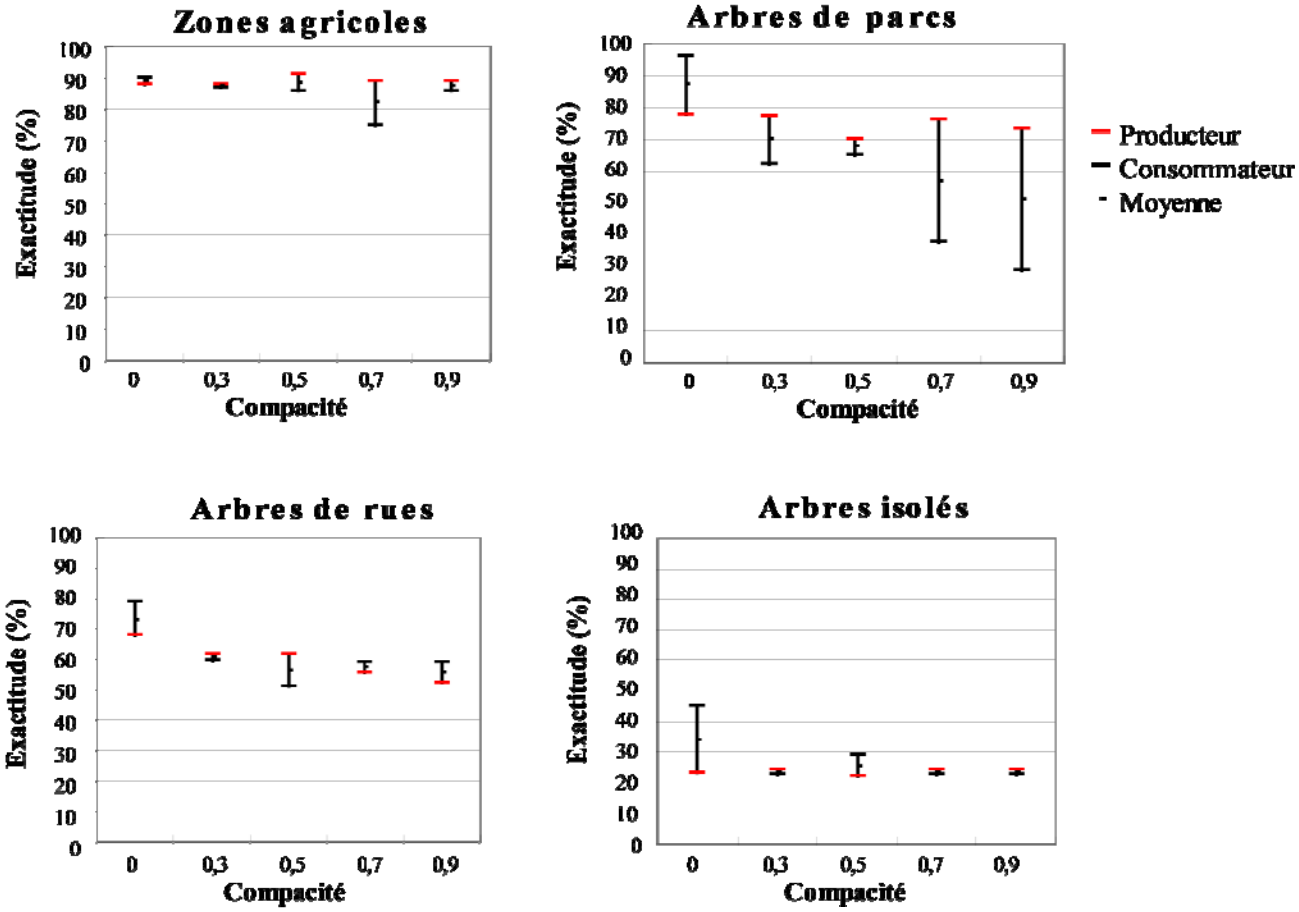


Figure 3. Exactitudes des segmentations pour quatre classes de végétation
Figure 3. Accucary of segmentations for 4 classes of vegetation

Tableau 4. Exactitude obtenue avec 4 bandes originales et rapport de couleur/forme = 0,7/0,3 de la classe de « végétation » (unité : %)
Table 4. Accuracy obtained with 4 original bands and colour/shape ratio = 0.7/0.3 of the « vegetation » class (unit : %)

Compacité	Arbre de parc	Arbre de rue
0	0,87	0,67
0,3	0,88	0,68
0,5	0,87	0,69
0,7	0,87	0,68
0,9	0,87	0,68

Pour trouver la meilleure valeur du ratio de compacité/lissage, nous avons effectué des tests en fixant le ratio de couleur/forme à 1/0 et en faisant varier la compacité de 0 à 0,9. Les niveaux d'exactitude de chaque classe donnés dans la figure 3 baissent quand la compacité augmente.

Pour expliquer cette baisse, nous avons comparé les segmentations aux niveaux de compacité 0,3 et 0,9 (dans les panneaux gauche et droite sur la figure 4) et nous avons trouvé une sur-segmentation au niveau de compacité de 0,9. Les meilleures exactitudes se trouvent dans la segmentation produite avec une compacité de 0 (Tableau 5).

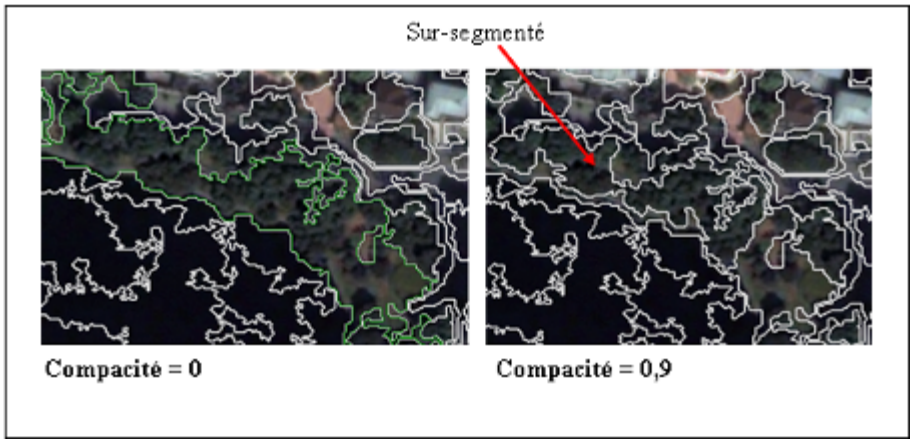


Figure 4. Segmentations d'une zone comprenant un parc (à 2 niveaux de compacité)
Figure 4. Segmentations for an area having a park (at 2 values of compactness)

Tableau 5. Exactitude de la meilleure segmentation des classes de végétation
Table 5. Accucary of the best segmentations for the classes of vegetation

Exactitude (%)	Arbres de parc	Arbres de rue	Arbres isolés	Zones agricoles	Toutes les classes
Producteur	0,78	0,68	0,23	0,88	0,64
Consommateur	0,96	0,79	0,45	0,90	0,78
Moyenne	0,87	0,73	0,34	0,89	0,71

En réalisant le processus de classification sur toute l'image, nous avons remarqué que dans certains endroits hors de la zone de test, des parties de parcs se sont connectées à des arbres de rues. Il est impossible de séparer ces deux parties par la segmentation. Nous avons donc découpé ces segments en deux parties à l'aide de l'outil manuel du logiciel.

Selon la validation par régions pré-sélectionnées, l'exactitude moyenne (la moyenne entre l'exactitude du producteur et du consommateur) des classes de végétation est de 89 %, 87 %, 73 % et 37 % pour les classes « zones agricoles », «arbres de parc », « arbres de rue » et « arbres isolés », respectivement. La faible exactitude obtenues pour les arbres isolés s'explique par le fait que sur le site de validation, il y a certains arbres individuels que l'œil ne peut pas détecter.

Les arbres isolés occupent seulement 0,4 % de la superficie de végétation dans le site de validation, ce qui est négligeable.

Selon la deuxième validation, les exactitudes du consommateur et du producteur pour la classe « zones agricoles » sont de 96% et 81%. Le fait que l'exactitude de production est plus faible s'explique par le fait qu'un certain nombre de points de référence ne sont pas détectés par la classification. Quant à la classe « arbres de rue », l'exactitude du consommateur est assez faible, soit 65%. L'erreur de commission élevée de cette dernière, est due au fait que la classification n'a pas réussi à séparer les groupes d'arbres de rue des groupes d'arbres de parcs. L'exactitude du producteur de cette classe est de 84%, ce qui acceptable en télédétection. L'erreur d'omission est due à des points correspondant aux arbres isolés et aux arbres bordant les îlots de bâtiments et les routes. Les niveaux exactitude du producteur des « arbres isolés » et des « arbres de parcs » sont très faibles, soit 35% et 47%, respectivement. Cela confirme de nouveau la difficulté de séparer les trois classes d'arbres. Nous croyons que pour bien séparer les arbres dans les rues et ceux dans les parcs, il faudrait avoir une carte d'utilisation du sol qui indique les trottoirs et les parcs.

3.3. Extraction de la classe « route »

Nous avons d'abord testé l'extraction de la seule grande route de la zone d'étude. La segmentation pour la classe « route » a été basée sur les quatre bandes d'origine de l'image Quickbird. Nous avons choisi différentes échelles pour décomposer l'autoroute en différents types d'objets en fonction de leur taille, de leur homogénéité et de la complexité de leur contour (Figure 5). La route se compose d'abord de grands polygones assez homogènes (rangés dans les classes « Route 1 » et « Route 2 » et différant par leur couleur). Elle contient aussi d'autres polygones qui sont moins homogènes à cause des objets correspondant à des véhicules et des bâtiments se trouvant sur ou au bord de la route. Ces polygones ont été rangés dans la classe « Route 3 ». Finalement, il y a de petits polygones au bord de la route qui ont un contour en zigzag à cause des petites maisons ou des arbres bordant la route. Ces polygones forment la classe « Route 4 ».

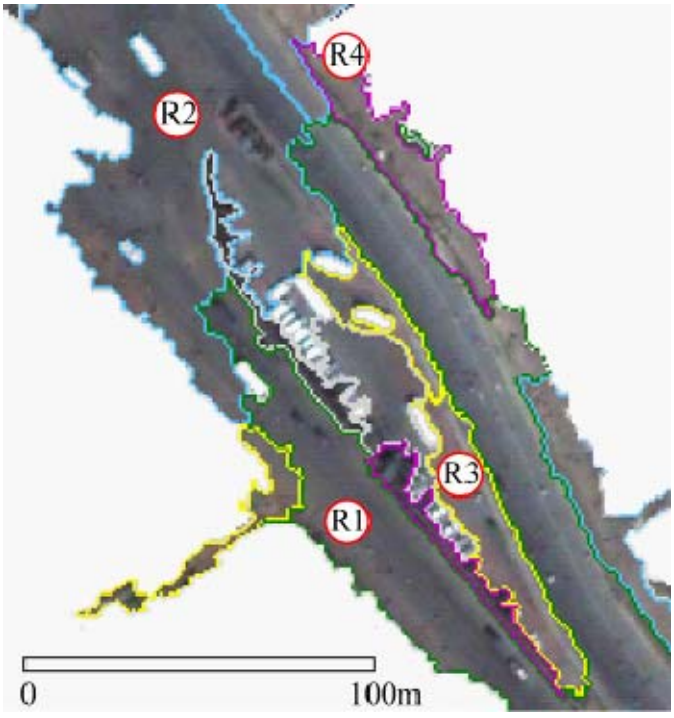


Figure 5. Quatre classes constituant de la route
Figure 5. Four sub-classes of the highway

Les valeurs des échelles ont été établies par une série de tests empiriques sur la segmentation de l'image. Nous avons testé cinq possibilités d'ordre de segmentation : 10 et 50, 15 et 50, 25 et 50, 35 et 50, et finalement 25, 35 et 50. Cette séquence de segmentation permet de garder le contour zigzag mais aussi de conserver les grands segments homogènes de la route. Lors de ces tests, nous avons fixé les paramètres de couleur à 0,8 et de la compacité à 0,3 parce que sur l'image les objets de type de route ont une couleur assez distincte et un contour lisse. L'exactitude de ces tests est illustrée dans la figure 6: la segmentation à l'échelle 50 basée sur une segmentation précédente à une échelle de 10 (appelé dorénavant 10-50) atteint le meilleur niveau d'exactitude et la segmentation à une échelle de 50 basée sur une segmentation précédente à une échelle de 15 (appelé 15-50) a le niveau le plus faible. L'exactitude de ces tests varie

fortement en fonction de l'échelle de segmentation. Cela suggère que les ratios de couleur/forme et de compacité/lissage ont une certaine influence sur l'échelle de segmentation. Nous avons donc testé d'autres segmentations avec des ratios différents pour les deux ordres de segmentation 15-50 et 10-50 (qui donnent le pire et le meilleur niveau d'exactitude avec le ratio de départ).

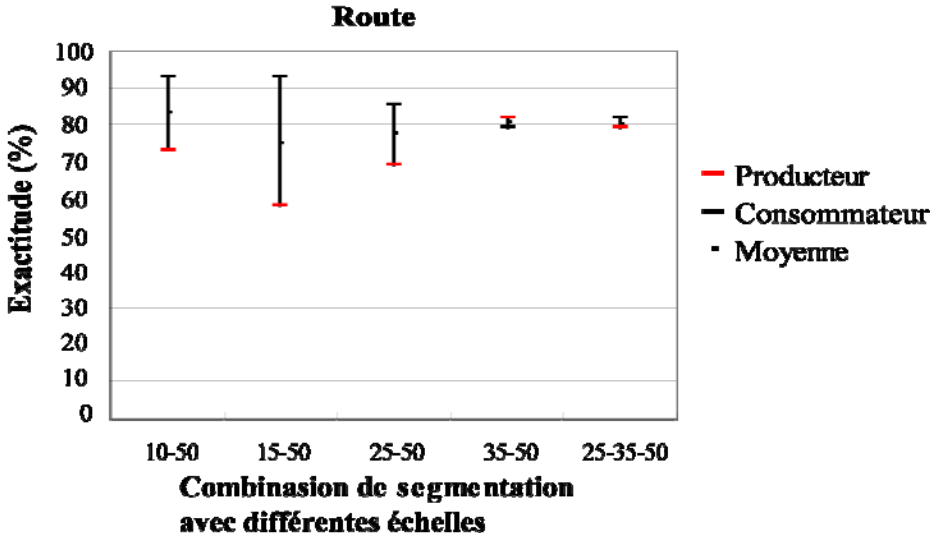


Figure 6. Exactitude de la classification de la route (à différents ordres de segmentations)
Figure 6. Accuracy of the classification of the highway (at different orders of segmentations)

Nous avons commencé par la segmentation de 15-50 avec différents ratios de couleur/forme, soit à 0,5/0,5 et 0,8/0,2, considérant que la couleur joue un rôle égal ou plus important que la forme dans l'identification de la route. À chaque niveau, nous avons utilisé trois ratios de compacité/lissage : 0/1, 0,3/0,7 et 0,5/0,5 (Figure 7). Les trois premières valeurs à gauche de la figure correspondent aux segmentations produites par l'indice de la couleur à 0,5 tandis que les trois valeurs à droite correspondent aux segmentations produites par l'indice de couleur à 0,8.

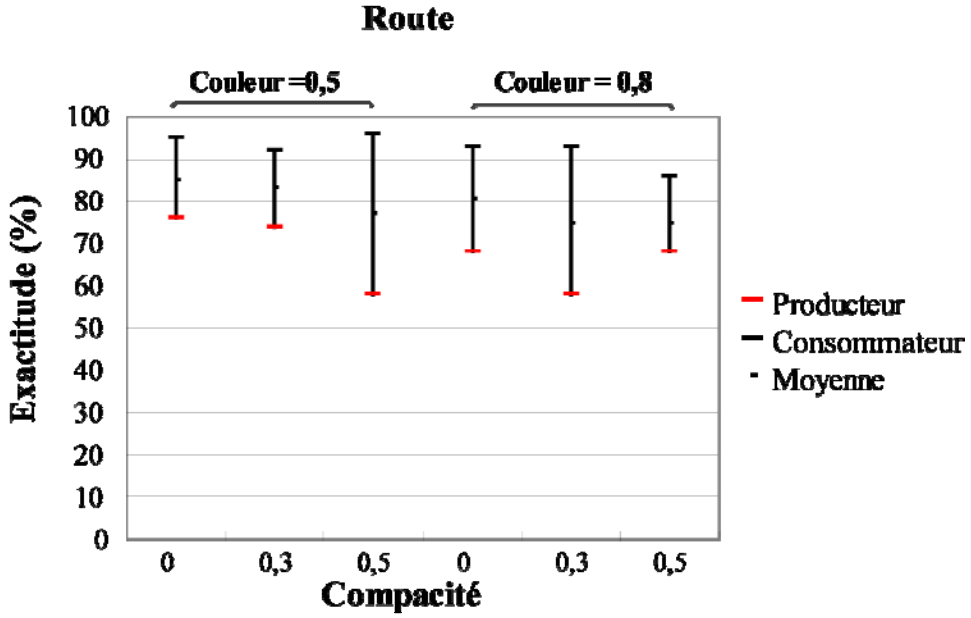


Figure 7. Exactitude de la classification de la route (échelle de 15-50)
Figure 7. Accuracy of the classification of the highway (scale of 15-50)

La tendance générale est que l'exactitude baisse quand la compacité des objets augmente. Cela s'explique par le fait que lorsque la compacité est privilégiée, les objets segmentés correspondant à la route sont moins allongés et regroupent donc parfois d'autres types de couverture des sols que la route, comme par exemple des bâtiments (Figure 8). La séparation de la route et les îlots de bâtiments devient plus difficile parce que ces deux classes ont une réponse spectrale très semblable. Des études précédentes ont déjà constaté de telle similarité entre les deux classes (Caprioli and Tarantino 2003; Puissant *et al.* 2006; Taubenböck *et al.* 2006). Il est à mentionner qu'à cause de la similarité spectrale

entre les surfaces de rues et les toits de bâtiments à Hanoi, la classification des petites rues n'était pas satisfaisante. Nous avons donc décidé de ne pas d'évaluer l'exactitude de la classification des petites routes.

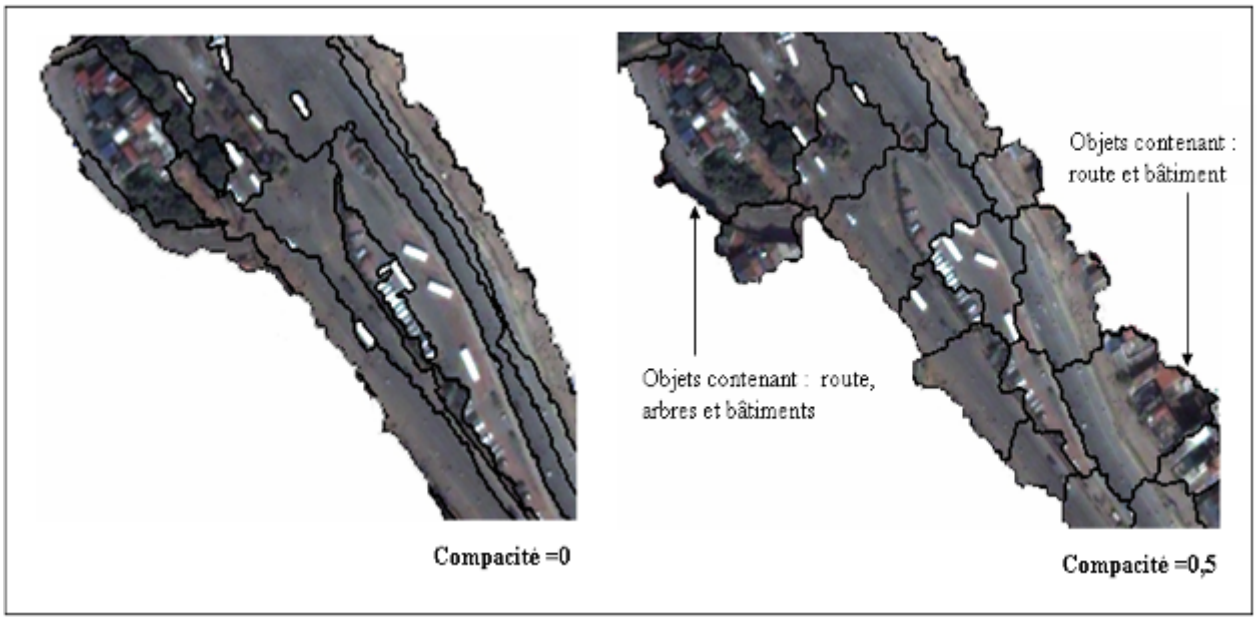


Figure 8. Segmentations de la route à 2 niveaux de compacité (échelle de 15-50 et couleur à 0,8)
Figure 8. Segmentations of the highway at 2 values of compactness (scale of 15-50 and colour of 0.8)

Ultimement, les tests des ratios de couleur/forme et de compacité/lissage appliqués à la segmentation à l'échelle 15-50 montrent que la meilleure exactitude est atteinte avec un ratio de couleur/lissage de 0,5/0,5 combiné avec un ratio de compacité/lissage de 0,5/0,5. Quant aux segmentations à l'échelle 10-50, leur exactitude suit la même tendance que ceux produits à l'échelle 15-50, mais ils sont moins variés (Figure 9). Le meilleur niveau d'exactitude moyenne à l'échelle 10-50, soit 84 %, est obtenu par la segmentation produite avec un ratio de couleur/forme de 0,8/0,2 et un ratio de compacité/lissage de 0,3/0,7.

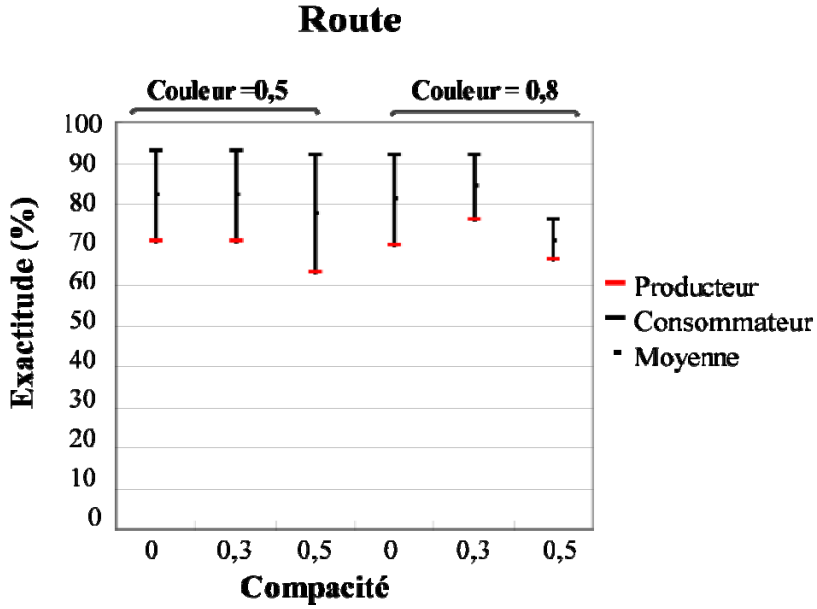


Figure 9. Exactitude de la classification de la route (échelle de 10-50)
Figure 9. Accuracy of the classification of the road (scale of 10-50)

En résumé, les niveaux d'exactitude de toutes les segmentations mentionnées dans cette section varient de 71 % à 85 %. Peu importe l'échelle et la couleur, les meilleures exactitudes (84 et 85 %) se trouvent quand la compacité est faible (0 ou 0,3). Cette dernière semble donc jouer un rôle essentiel dans la segmentation de la route et il est préférable que cette valeur soit minimale. Selon la deuxième validation par point, l'exactitude du consommateur et du producteur sont de 91% et 94%, ce qui est très satisfaisant. La matrice de confusion révèle cependant que certains arbres sont confondus dans la classe de « route » et que certaines parties de la route ont été omises.

3.4. Extraction de la classe « îlots résidentiels »

Cette classification est basée sur les classes de végétation et sur le réseau routier qui provient de sources cartographiques à l'échelle de 1 : 50 000 établies en 2005. Nous avons créé une image binaire composée de 0 et de 1, dont les pixels associés à la végétation et au réseau routier ont la valeur 0. Nous avons vectorisé cette bande (avec le logiciel ArcGIS) afin d'obtenir les polygones des zones de végétation et de routes. La segmentation a été ensuite effectuée sur une combinaison de quatre bandes d'origine, de la bande raster et de la bande vectorielle. Après plusieurs essais de segmentation, nous avons fixé l'échelle à 70. Nous avons fixé le ratio de compacité/lissage à 0,9/0,1 et varié la valeur de couleur de 0,1 à 0,8. Cette forte valeur de compacité a été retenue parce que les îlots ont une forme compacte, carrée ou rectangulaire. Les niveaux d'exactitudes changent légèrement en fonction de la couleur et ils ont tendance à augmenter dans la même direction (Figure 10). La meilleure exactitude est obtenue quand la valeur de couleur est à 0,5.

La meilleure segmentation pour les bâtiments résidentiels atteint 80 % d'exactitude avec un indice de couleur de 0,5 et un indice de compacité de 0,6. Ces valeurs peuvent s'expliquer par le fait que les îlots ont une forme particulière, le paramètre de couleur jouant un rôle moins important lors de la segmentation des îlots que des autres classes. Le paramètre de compacité, quant à lui, résulte en une sur-segmentation s'il est bas (le panneau gauche de la figure 12), mais en une segmentation adéquate s'il a une valeur élevée (le panneau droit de la figure 12). Ceci s'explique par le caractère compact des formes rectangulaires associées aux îlots résidentiels. Selon la deuxième validation, l'exactitude du consommateur et du producteur sont de 90% et 95%, ce qui est très satisfaisant. Comme dans la classification de « route », nous observons que certains arbres ont été confondus dans la classe « îlots résidentiels » et que certaines parties des îlots n'ont pas été détectées par la classification.

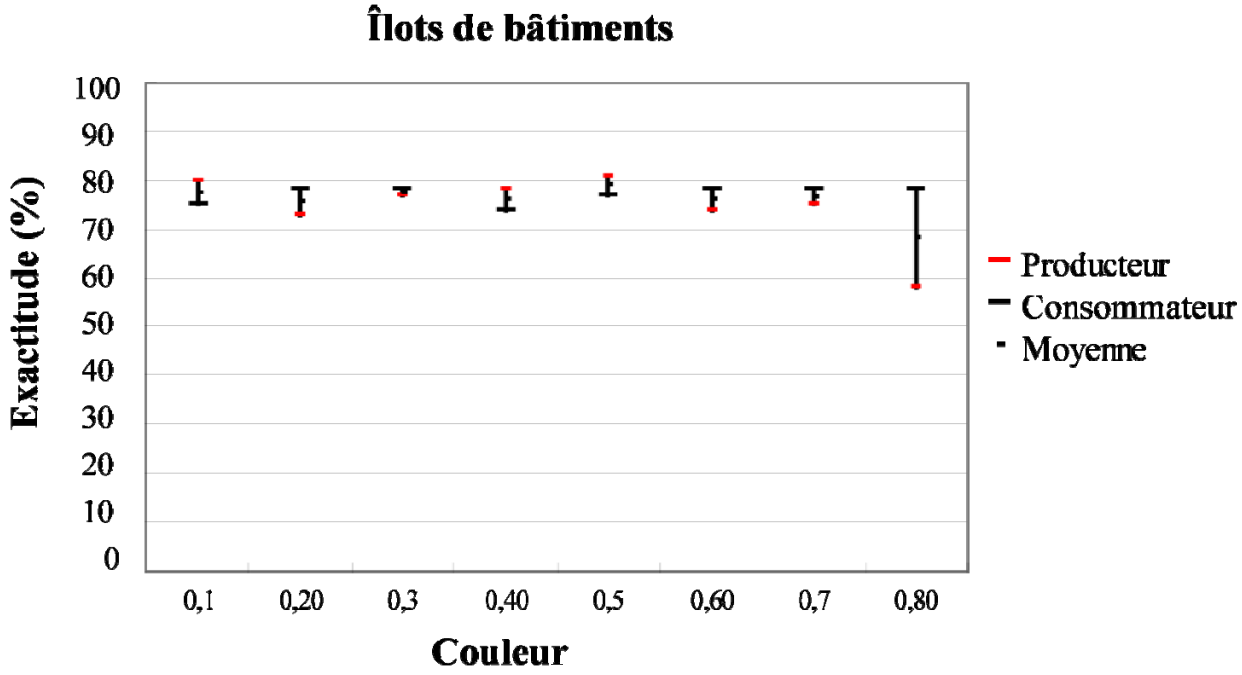


Figure 10. Exactitude des segmentations de « îlots de bâtiments » (compacité à 0,9)
Figure 10. Accuracy of segmentations of « housing blocks » (compactness of 0.9)

En fixant la couleur à 0,5 nous avons testé des valeurs de compacité variant de 0,1 à 0,9 (Figure 11). Les niveaux d'exactitude ont tendance à augmenter pour se stabiliser à partir d'une valeur de compacité de 0,6.

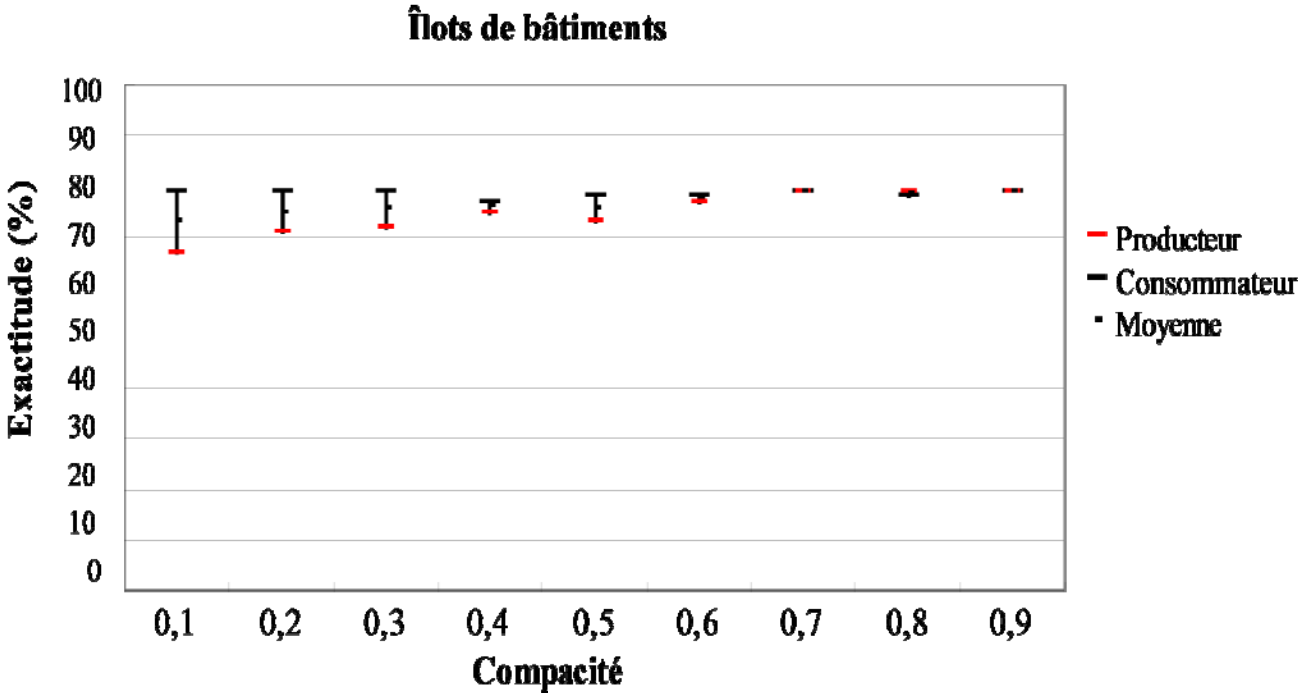


Figure 11. Exactitude des segmentations de « îlots de bâtiments » (couleur à 0,5)
Figure 11. Accuracy of segmentations of « housing blocks » (colour of 0.5)

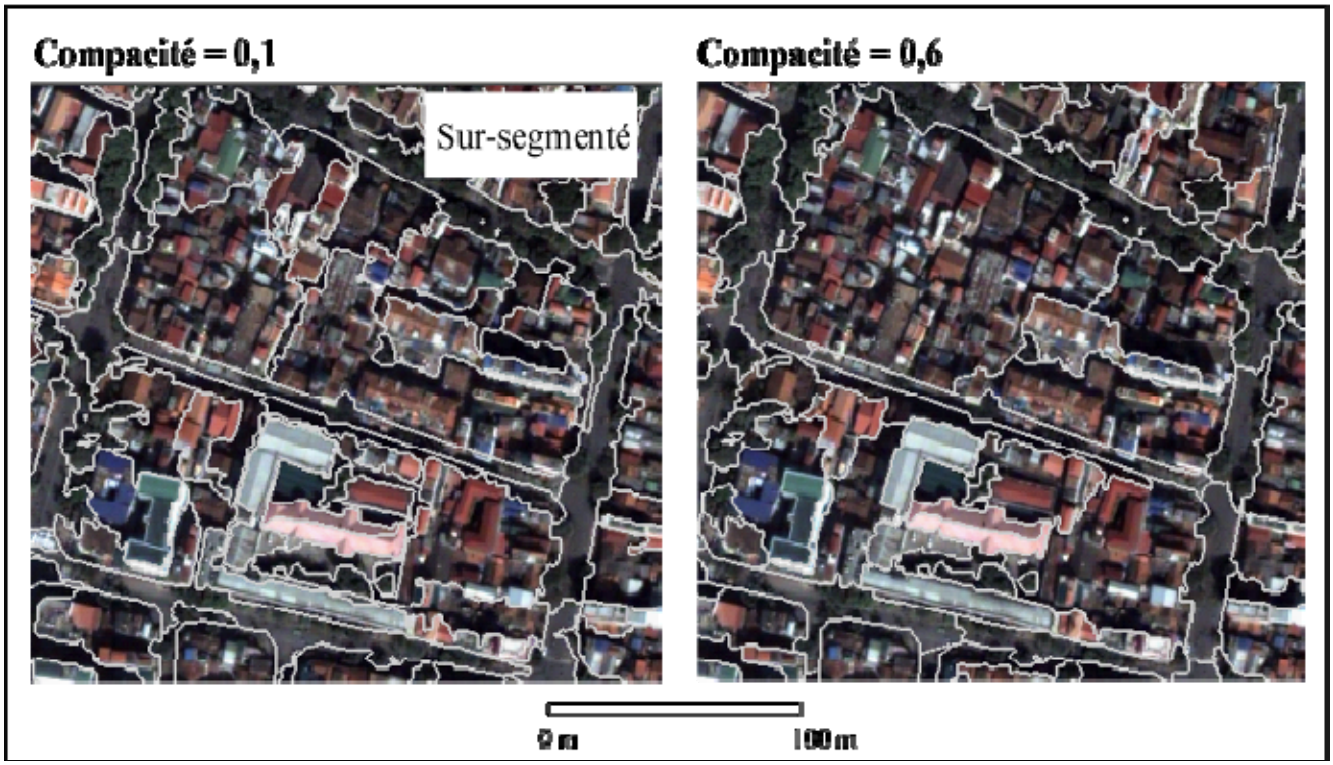


Figure 12. Comparaison des segmentations de « îlots de bâtiments » à deux niveaux de compacité
Figure 12. Comparation of segmentations for « housing blocks » at 2 values of compactness

Nous avons superposé toutes les classes obtenues dans une carte finale (figure 14). Les chevauchements entre les classes sont négligeables (0,4 %) et ne sont pas visuellement détectables. Quant à la segmentation, les paramètres finaux sont résumés dans le Tableau 6.



Figure 13. Classification finale de l'image Quickbird
Figure 13. Final classification of the Quickbird image

Tableau 6. Meilleurs paramètres de segmentations et exactitudes de toutes les classes
Table 6. Best parameters of segmentations and accuracy of all classes

Classes	**Exactitude moyenne (%)**	**Meilleures valeurs des paramètres de segmentation**			
		Échelle	Bandes	Couleur	Compacité
Lacs	97	100	4 bandes Quickbird	1	0
Fleuves	98	100	4 bandes Quickbird	1	0
Zones agricoles	91	100	NDVI	0,3	0,5
Arbres de parc	87	50	NDVI	1	0
Arbres de rue	73	25	NDVI	1	0
Arbres isolés	34	25	NDVI	1	0
Routes	84	10-50	4 bandes Quickbird	0,8	0,3
	85	15-50		0,5	0
Ilots résidentiels	80	70	4 bandes Quickbird	0,5	0,6

4. DISCUSSION

Notre méthode a permis d'extraire les huit classes avec une exactitude moyenne de 80,5 % selon la validation ciblant certaines zones, et 78% selon la validation par points sur l'image entière. Le niveau d'exactitude le plus élevé est obtenu pour la classe « fleuve » et le plus faible pour la classe « arbre isolé ».

La catégorie des plans d'eau (comprenant les lacs et le fleuve) atteint un niveau d'exactitude très élevé, ce qui est dû à la réponse spectrale très basse de l'eau et qui est très distincte des autres classes. Ce fait est confirmé par d'autres travaux de traitement de l'image par l'approche d'objet (Flanders *et al.*, 2003; Lang *et al.*, 2006; Platt and Rapoza, 2008).

Dans la catégorie de végétation, nous avons obtenu des résultats adéquats, sauf pour la classe « arbres isolés ». La faible exactitude de cette classe s'explique par le fait qu'il est difficile d'identifier les arbres dispersés, même visuellement. Laliberte *et al.* (2004) confirment aussi la difficulté d'identifier les arbustes de petite taille dans leur classification. Plusieurs études récentes ont suggéré que les données du LiDAR pourraient être utiles pour identifier les arbres dispersés en milieu urbain (par exemple Sung 2012; Zhou and Troy 2008).

Les paramètres de segmentation proposés pour la végétation par les auteurs précédents privilégient la couleur par rapport à la forme, mais n'attribuent aucun poids (Laliberte *et al.*, 2004) ou seulement un poids minime à la compacité (Lang *et al.*, 2006; Platt and Rapoza, 2008). Nos résultats corroborent bien avec ces constats en montrant que que les meilleurs niveaux d'exactitude sont obtenus dans les segmentations avec des valeurs de couleur de 1 et de compacité de 0. Cependant, à Hanoi, il n'y a pas de grands parcs régionaux avec des groupes d'arbres à taille importante. Dans des villes où il y a de tels types de parcs, nous recommandons de faire preuve de prudence lors du choix d'échelle et de compacité de segmentation, qui peut avoir un impact sur la taille et la forme des objets segmentés.

Pour la classe « route » nous avons réussi à extraire l'unique grande route de la zone d'étude avec une exactitude de 85 %. L'extraction de cette catégorie est très complexe tant au niveau de la segmentation qu'au niveau de l'utilisation des fonctions d'appartenance de la classification. Nos tests de segmentation démontrent que la couleur ne manifeste pas d'influence sur la variabilité d'exactitude mais qu'une valeur supérieur à 0,5 donne une meilleure exactitude. L'influence de la compacité est plus nette; il est préférable qu'elle soit faible (de 0 à 0,3) car la classe a une forme allongée (incidemment, deux combinaisons distinctes de paramètres de segmentation nous permettent d'atteindre la même exactitude). Quant aux fonctions d'appartenance, nous avons utilisé 11 règles à cause de la difficulté de séparer la route et des bâtiments l'entourant. Nous constatons d'importantes confusions entre les deux catégories, dues à leur similarité spectrale sur l'image Quickbird. La classe de route est également jugée difficile à extraire par Puissant *et al.* (2006), Taubenböck *et al.* (2006) et Caprioli and Tarantino (2003). C'est la raison pour laquelle les petites rues de la zone d'étude ne sont pas adéquatement segmentées. Dans les villes densément construites comme Hanoi où les rues de

petites tailles sont répandues, nous recommandons d'utiliser des images à meilleures résolutions spatiales et spectrales, ainsi que des méthodes de segmentation plus adaptées aux routes (voir par exemple Zhang and Couloigner 2006).

Pour la classe d'îlots de bâtiment, le niveau d'exactitude obtenu est de 80 %. Les tests de segmentation révèlent que la forme est plus pertinente pour les îlots de bâtiments (0,5) que pour les segmentations d'autres catégories car ils ont une forme particulière. De plus, une forte compacité (0,6) est utile pour segmenter les objets ayant une forme rectangulaire compacte comme ceux de cette classe. Nos paramètres optimaux sont légèrement plus faibles que ceux proposés par Flanders *et al.* (2003) pour des zones de découpages forestières (forme = 0,8 et compacité = 0,9). Les îlots de bâtiments à Hanoi sont assez variés au niveau de leur réponse spectrale et de leur texture, ce qui est le cas des villes dans les pays en développement. Il nous a donc fallu huit règles de classification pour arriver à bien extraire cette classe. Dans les villes moins denses et mieux planifiées, il est possible d'identifier des bâtiments individuels (voir par exemple Sebari and He 2009)

En accord avec les autres auteurs, nos résultats démontrent que les paramètres de segmentation sont plus importants que les règles de classification. La segmentation joue donc un rôle primordial dans la qualité de l'approche orientée objet, comme confirment les auteurs précédents (Dey *et al.* 2010; Taubenböck *et al.* 2006).

5. CONCLUSIONS

L'arrivée des images de très haute résolution spatiale offre une nouvelle source d'information pour les applications en milieu urbain. Dans cet article, nous avons effectué une procédure de l'extraction des objets urbains à partir d'une image Quickbird par une approche orientée objet. Malgré la faible résolution spectrale de l'image, cette approche donne les résultats encourageants pour presque tous les objets d'intérêts, à l'exception des classes de « arbres isolées » et des petites rues. Nous sommes arrivés à des meilleures valeurs pour les paramètres de segmentation, qui n'avaient jamais été testés de manière empirique et systématique. Les valeurs proposées correspondent aux caractéristiques de couleur et de forme des objets dans une ville dense comme Hanoi. Elles pourraient être des suggestions initiales pour des tests de segmentations dans les futurs travaux de l'extraction des objets urbains à partir d'une image à THRS.

Pour chaque classe, nous avons aussi établi une stratégie de tests pour les paramètres de segmentation afin de choisir leurs meilleures valeurs de manière moins intuitive et subjective. Cette démarche exige plusieurs séquences de tests, mais avec de récents développements permettant l'utilisation du logiciel eCognition pour le traitement automatique de plusieurs images à la fois, en font une solution efficace et fiable pour les applications de l'approche orientée objet.

Les résultats obtenus dans cet article permettent de conclure que la télédétection à haute résolution spatiale est très utile pour des mise-à-jours et inventaires des bases de données à faible coût et dans des délais rapides. L'utilisation de l'approche orientée en combinaison avec ces nouvelles images est particulièrement adaptée au contexte des pays en développement où l'urbanisation se déroule à grande vitesse et où les bases de données sont incomplètes ou difficilement accessibles.

6. Remerciements

Nous remercions les centres de recherche VTGEO de l'Institut de géologie (Académie des sciences du Viêt-Nam) et CARGIS de l'Université des sciences de Hanoi qui nous ont fourni l'image Quickbird et les cartes de l'infrastructure de Hanoi. Notre étude a été financée par le programme de bourse de l'Agence universitaire de la Francophonie et des subventions de recherche du Conseil de recherches en sciences naturelles et en génie du Canada.

7. Références

Baatz, M., Benz, U. C., Dehshani, S., Heynen, M., Höltje, A., Hofmann, P., Lingenfelder, I., Mimler, M., Sohlbach, M., Weber, M., Willhauck, G., GmbH, D. I., 2004, eCognition Professional, Userguide version 4.0, München, pp. 475.

Benz, U. C., Hofmann, P., Willhauck, G., Lingenfelder, I., Heynen, M., 2004, Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information, ISPRS Journal of Photogrammetry & Remote Sensing 58:239-258.

Bhaskaran, S., Paramananda, S., Ramnarayan, M., 2010, Per-pixel and object-oriented classification methods for mapping urban features using Ikonos satellite data, Applied Geography 30:650-665.

Bonn, F., Rochon, G., 1992, Précis Teledetection, Volume 1 : Principes et méthodes, Presses de l'Université du Québec, Québec, pp. 485.

Caprioli, M., Tarantino, E., 2003, Urban features recognition from VHR satellite data with an object-oriented approach, ISPRS, Challenges in Geospatial Analysis, Integration and Visualization II, Stuttgart.

Carleer, A. P., Debeir, O., Wolff, E., 2005, Assessment of very high spatial resolution satellite image segmentations, Photogrammetric Engineering & Remote Sensing Vol. 71(11):1285–1294.

Cheng, H. D., Jiang, X. H., Wang, J., 2001, Color image segmentation: advances and prospects, Pattern Recognition 34:2259-2281.

Congalton, R. G., Green, K., 2009, Assessing the Accuracy of Remotely Sensed Data—Principles and Practices (Second edition), CRC Press, Taylor & Francis Group, Boca Raton, FL, pp. 183.

Coulibaly, L., Gwyn, Q. H. J., 2005, Intégration de données d'images satellitaires optiques et radars et de données topographiques pour la cartographie géomorphologique, Journal canadien de télédétection 31(6):439–449.

Dang, X. D., Le, H. K., 2004, La population de Hanoi, in: Gubry, P. Eds., Population et développement au Viêt-nam. La maison d'édition du monde, Hanoi, pp. 243-262.

Dey, V., Zhang, Y., Zhong, M., 2010, A review on image segmentation techniques with remote sensing perspective, ISPRS TC VII Symposium – 100 Years ISPRS (Wagner W., Székely, B., eds.), Vienne, pp. 31-42.

Drakakis-Smith, D., Kilgour, A., 2001, Sustainable urbanization and environmental issues in Vietnam, in: Adger, N. W., Nelly, M. P., Nguyen, H. N. (Eds.), Living with environmental change: Social vulnerability, adaption and resilience in Vietnam. Taylor and Francis Group, London and New York, pp. 213-233.

Flanders, D., Hall-Beyer, M., Pereverzoff, J., 2003, Preliminary evaluation of eCognition object-based software for cut block delineation and feature extraction, Canadian Journal of Remote Sensing 29(4):441–452.

Guindon, B., 1997, Computer-Based aerial image understanding: A review and assessment of its application to planimetric information extraction from very high resolution satellite images, Canadian Journal of remote Sensing 23(1):38-47.

Hay, G. J., Castilla, G., 2006, Object-based image analysis: strengths, weaknesses, opportunities and threats (SWOT), ISPRS 1st International Conference on Object-based Image Analysis (OBIA 2006), Salzburg.

Herold, M., Gardner, M., Hadley, B., Roberts, D., 2002, The spectral dimension in urban land cover mapping from high resolution optical remote sensing data, 3rd Symposium on Remote Sensing of Urban Areas, Istanbul, pp. 8.

Herold, M., Gardner, M. E., Roberts, D. A., 2003, Spectral resolution requirements for mapping urban areas, IEEE transactions on geoscience and remote sensing 41(9):1907-1919.

Laliberte, A. S., Rango, A., Havstad, K. M., Paris, J. F., Beck, R. F., McNeely, R., Gonzalez, A. L., 2004, Object-oriented image analysis for mapping shrub encroachment from 1937 to 2003 in southern New Mexico, Remote Sensing of Environment 93:198-210.

Lang, S., Jekel, T., Hölbling, D., Schöpfer, E., Prinz, T., Kloyber, E., Blaschke, T., 2006, Where the grass is greener - mapping of urban green structures according to relative importance in the eyes of the citizens, First Workshop of the EARSeL - Special Interest Group on Urban Remote Sensing, Berlin.

Longley, P. A., Barnsley, M. J., Donnay, J. P., 2001, Remote sensing and urban analysis: a research agenda, in, Remote sensing and urban analysis. Taylor and Francis, London and New York, pp. 268-268.

Miller, R. B., Small, C., 2003, Cities from space: potential applications of remote sensing in urban environmental research and policy, Environmental Science & Policy 6:129-137.

O'Rourke, D., 2005, Community-driven regulation: balancing development and the environment in Vietnam, The MIT Press, Cambridge, pp. 297.

Platt, R. V., Rapoza, L., 2008, An evaluation of an object-oriented paradigm for land use/land cover classification, Professional Geographer 60:87-100.

Procacci, F., Luong, T. T., 2007, Learning from KTT: Analysis of the collective residential quarters of Hanoi, International Conference on Sustainable Architectural Design and Urban Planning, Hanoi Architectural University, Hanoi.

Puissant, A., Weber, C., Sheeren, D., Wemmert, C., Gançarski, P. F., 2006, Amélioration des connaissances sur l'environnement urbain : intérêt de l'intégration de règles dans les procédures de classifications (Enhancing understandings of urban environment: Interest of rule-integrated classifications), Interactions Nature-Société, analyse et modèles, La Baule, pp. CDrom [**in French**].

Sebari, I., He, D.-C., 2009, Approach to nonparametric cooperative multiband segmentation with adaptive threshold, Applied Optics 48(20):3967-3978.

Sugumaran, R., Voss, M., 2007, Object-oriented classification of LIDAR-fused hyperspectral imagery for tree species identification in an urban environment, 2007 Urban Remote Sensing Joint Event, Paris.

Sung, C. Y., 2012, Evaluating the efficacy of a local tree protection policy using LiDAR remote sensing data, Landscape and Urban Planning 104(1):19-25.

Taubenböck, H., Esch, T., Roth, A., 2006, An urban classification approach based on an object-oriented analysis of high resolution satellite imagery for a spatial structuring within urban areas, 1st EARSeL Workshop of the SIG Urban Remote Sensing Berlin, Germany.

Thomas, N., Hendrix, C., Congalton, R. G., 2003, A comparison of urban mapping methods using high-resolution digital imagery, Photogrammetric Engineering & Remote Sensing 69(9):963–972.

Tô, K., 2008, Conservation pressing task and new documentation of old tube houses in Hanoi old quarter through the case of no. 47 Hang Bac street house, Journal of Architecture and Planning 73(624):457-462.

United Nations, 2008, World urbanization prospects: The 2007 revision highlights.

Wuest, B., Zhang, Y., 2009, Region based segmentation of QuickBird multispectral imagery through band ratios and fuzzy comparison, ISPRS Journal of Photogrammetry and Remote Sensing 64(1):55-64.

Yu, Q., Gong, P., Clinton, N., Biging, G., Kelly, M., Schirokauer, D., 2006, Object-based detailed vegetation classification with airborne high spatial resolution remote sensing imagery, Photogrammetric Engineering & Remote Sensing 72(7):799-811.

Yuan, F., Bauer, M. E., 2003, Mapping impervious surface area using high resolution imagery: a comparison of object-based and per pixel classification, ASPRS Annual Conference, Reno, Nevada, pp. 8.

Zhang, Q., Couloigner, I., 2006, Automated road network extraction from high resolution multi-spectral imagery, ASPRS 2006 Annual Conference, Reno, Nevada.

Zhang, Y. J., 1997, Evaluation and comparison of different segmentation algorithms, Pattern Recognition Letters 18:963-974.

Zhang, Y. J., 2006, Chapter 1: An overview of image and video segmentation in the last 40 years, in Advances in image and video segmentation. IRM Press, Hershey, PA, pp. 1-15.

Zhou, W., Troy, A., 2008, An object-oriented approach for analysing and characterizing urban landscape at the parcel level, International Journal of Remote Sensing 11(10):3119-3135.