



EVALUATION QUANTITATIVE DE LA QUALITE DE SEGMENTATION D'IMAGES DE TELEDETECTION

Nicolas PAPARODITIS ^{1,2}- Michel CHAPRON ²- Jean-Pierre COCQUEREZ ²

1 - AEROSPATIALE - Espace & défense - SE/TO/T - BP 99 - 06322 Cannes la Bocca Cedex

2 - ENSEA - Laboratoire ETIS - 6, avenue du Ponceau - 95014 Cergy-Pontoise Cedex

RESUME

Un problème crucial se pose dans l'exploitation de données de télédétection. Ce problème est l'évaluation des saisies ou primitives images, linéiques ou surfaciques, extraites par un opérateur manuel ou automatique pour certaines conditions données. Cette validation qualitative habituellement, est réalisée par une mesure quantitative de la qualité du produit extrait par rapport à une référence. En général, les saisies sont comparées à une saisie manuelle qui a ses propres erreurs et qui est non représentative de la réalité du paysage. On utilise le principe de simulation d'images pour comparer les saisies à une référence réaliste à la résolution de l'image.

INTRODUCTION

La majeure partie des informations cartographiques extraites des images de télédétection aériennes ou satellitaires sont issues de saisies manuelles monoscopiques ou stéréoscopiques réalisées par un photo-interprète. Son travail consiste à extraire des objets cartographiques symboliques linéiques ou surfaciques (bâti, ponts, routes, voies ferrées, cours d'eau, traits de côtes, etc) pour la réalisation de cartes thématiques. Ces restitutions enrichissent en général, des Bases de Données Géographiques ou des Systèmes d'Informations Géographiques, multi-utilisateurs où le besoin de mise à jour (les changements cartographiques ayant lieu, pour la grande majorité, sur des constructions ou destructions humaines) est aussi important que la connaissance de la qualité des données. Ces données étant destinées à des utilisateurs variés, les besoins de précision en terme de qualité sont eux aussi variés.

L'accroissement du besoin géographique et la multiplication des données de télédétection de plus en plus résolues, rendent ce travail de mise à jour de plus en plus difficile. C'est la raison pour laquelle la demande d'opérateurs automatiques d'extraction d'objets cartographiques se fait de plus en plus pressante. Les résultats de ces opérateurs automatiques sont validés en comparaison à une saisie manuelle, celle-ci étant alors supposée parfaite. C'est pour prendre en compte les imperfections de cette saisie manuelle de référence que nous proposons une méthode de validation quantitative s'appuyant sur la simulation d'images. Le modèle de paysage sur lequel s'effectue la simulation fournit une "vérité terrain" à laquelle on peut comparer l'information restituée. La difficulté d'une telle comparaison réside alors dans la définition d'une distance géométrique qui rende compte des similitudes morphologiques entre la saisie (segmentation) manuelle ou automatique et la carte de référence sachant que celle-ci est elle-même modélisée à une échelle donnée.

Nous commencerons par énoncer le problème de l'évaluation des méthodes de segmentation avant d'aborder la

ABSTRACT

The exploitation of remote sensing data raises a crucial problem which is the quality assessment of feature extraction techniques by an either manuel or automatic operator for a given set of conditions. This assessment which is usually qualitative is obtained by a quantitative measurement of the extracted product as compared to a given reference. The image features are often compared with an unprecise manual feature selection far from the ground truth. Image simulation is used to compare the feature extraction with an ideal reference upon which image simulations are based.

présentation de la méthode de comparaison entre segmentation et référence et l'apport des méthodes de simulation pour cette comparaison. On cherchera aussi à analyser l'influence des paramètres images sur la méthode de segmentation.

VALIDATION D'ALGORITHMES DE SEGMENTATION

La restitution d'objets cartographiques par des méthodes de segmentation bas-niveau, qu'elles soient contour ou région, sur des images de télédétection est un problème complexe. En effet, les images de télédétection optique se caractérisent par leur richesse texturale liée à la diversité géométrique et radiométrique des objets observés à travers les échelles. Ces textures provoquent des sur ou des sous segmentations locales des objets symboliques par rapport à une spécification de saisie manuelle. Ces problèmes de non respect du symbolisme sont provoqués par les propriétés locales du signal dans les images qui ne sont pas prises en compte par des algorithmes aux propriétés globales. Une revue complète des méthodes de segmentation et des problèmes associés est donnée dans [COC95].

Les opérateurs de segmentation sont nombreux et de qualité variable en fonction des caractéristiques des images. Quel opérateur choisir pour une application et un type d'images donné? La nécessité de l'estimation quantitative, relative et absolue, de la qualité des opérateurs se pose. Cette estimation soulève des problèmes complexes. Faut-il privilégier des opérateurs restituant une bonne symbolique cartographique plutôt qu'une bonne localisation des frontières ou les deux de manière pondérée? On donne à la segmentation le qualificatif de méthode bas-niveau alors qu'elle ne l'est pas réellement. En effet, les critères ou paramètres de la segmentation sont ajustés, pour une application concrète, par un opérateur humain qui sait que ceux-ci doivent être choisis pour que la ressemblance entre la segmentation automatique et sa propre segmentation de l'image soit optimale. La part d'interprétation ou plutôt de spécification dans une segmentation est donc importante. On



doit donc en tenir compte lors de l'évaluation de la qualité intrinsèque de cet algorithme.

Evaluer quantitativement le résultat de méthodes de segmentation en travaillant sur des images synthétiques de tests non réalistes n'est pas représentatif. Ces images permettent de donner une idée comportementale des algorithmes mais ne permettent pas d'évaluer leurs performances sur des images complexes comme les images aériennes ou satellitaires.

Par ailleurs, il est difficile, voire impossible, de valider un algorithme de segmentation à partir d'un petit nombre de tests, qui **valideraient alors le produit segmenté et non pas la méthode de segmentation**. En effet, chaque image est caractérisée par un paysage particulier, un capteur particulier, des conditions d'éclairage particulières, des conditions d'acquisitions particulières.

METHODES DE COMPARAISON A UNE REFERENCE

Une segmentation ne peut être validée qu'à partir de données de référence issues d'une source externe. Lorsque des données de référence sont disponibles et fiables, une validation externe est envisageable: elle consiste à mesurer l'écart entre la carte obtenue et la référence. Cette méthode, qui est la plus usuelle, se heurte cependant à deux types d'obstacles.

D'abord les données de référence sont peu souvent adaptées à ce type de validation, que ce soit parce qu'elles sont en nombre insuffisant (conduisant alors à des statistiques dénuées de sens) ou parce qu'elles sont elles mêmes entachées d'une erreur. Ensuite la définition de la distance entre la segmentation et la référence n'est pas triviale. En effet, elle doit être pertinente d'un point de vue statistique et refléter le type d'information que la segmentation est supposée fournir.

L'évaluation de la qualité de segmentation par rapport à une référence n'a de sens que si la référence est de meilleure qualité que la segmentation. On peut dire alors que la référence est plus proche de la réalité ou du terrain nominal. Une référence obtenue par saisie manuelle à l'échelle de l'image, ne permettra pas de réaliser une évaluation de la segmentation car elle est entachée d'erreurs qui même si elles sont subpixelaires sont trop importantes pour la précision de la mesure voulue. Ainsi la référence est en général issue de saisies manuelles d'images stéréoscopiques à des résolutions plus fines ou, issue de relevés topographiques 3D sur le terrain. Cette référence 3D, pour être comparée à la segmentation, doit alors être projetée dans l'espace image. Si la référence est issue de saisies image, les objets cartographiques sont extraits sous forme de points (caractéristiques ou non caractéristiques) ou de suites de segments s'intersectant en ces points. Ces ensembles de segments consécutifs peuvent être fermés pour des objets surfaciques ou ouverts dans le cas d'objets linéaires. Les segments saisis ont une précision subpixelaire qui n'a pas de limitation théorique, mais qui dépend du facteur de zoom, de l'interpolateur associé et de l'acuité stéréoscopique de l'opérateur.

La segmentation est réalisée sur une image contenant des objets cartographiques de nature spatio-fréquentielle différente tant d'un point de vue radiométrique que géométrique. La segmentation s'apparentant à un filtrage spatial de l'information, on peut penser avec certitude que la qualité de leur segmentation ne sera pas homogène. Il semble donc important de les séparer en différentes classes fréquentielles et de mesurer la qualité de la segmentation pour ces différentes classes. Les partitions en classes pourront se faire de la manière suivante:

- par type de représentation: ponctuel, linéaire, surfacique;
- par type sémantique: bâti, ponts, rivières, routes, etc;
- par taille: (pour le bâti) industriel, collectif, pavillonnaire.

Pour chacune de ces classes, la comparaison à une référence peut être envisagée: par contrôle ponctuel, et par contrôle linéaire. Ces deux types de comparaison se heurtent à des contraintes.

Contrôle ponctuel

Cette méthode consiste à évaluer la qualité de la segmentation à partir de points homologues connus à la fois dans la segmentation et dans la référence. La qualité de l'estimation de points homologues se mesure par la distance entre deux points homologues (distance euclidienne par exemple). Pour un ensemble de points homologues, l'estimation globale de la qualité est donnée en général par l'erreur moyenne quadratique (emq) si la référence est exacte ou l'écart moyen quadratique (ecmq) sinon. [ABB94] donne une description détaillée du contrôle ponctuel.

La connaissance de points homologues est ambiguë car la segmentation ne repose pas sur la notion de points caractéristiques cartographiques. La recherche même manuelle du meilleur homologue possible pour un point de la référence dans l'ensemble des points frontières de l'image segmentée serait approximative. Cette comparaison ne refléterait pas la qualité globale de la segmentation.

Contrôle linéaire

La comparaison entre la segmentation et la référence se fait maintenant par comparaison de l'ensemble des points de couples de contours homologues. Hausdorff introduit une distance entre deux objets qui traduit leur ressemblance topologique. Cette distance est très sensible au bruit sur l'estimation des contours, par la présence dans sa définition des opérateurs de bases de la morphologie mathématique (sup. et inf.). Elle mesure la plus grande des plus distances de tous les points d'un contour à l'autre contour. Afin d'éviter de mesurer une distance maximale entre les deux contours, [BAD92] propose une généralisation de cette distance qui va donner une distance moyenne entre les deux contours. [HOT95] s'affranchit de ces bruits, avec une méthode fondée sur la simulation, par moyennage des distances de Hausdorff calculées sur des nouveaux contours obtenus par bruitage des originels.

Problèmes liés à l'échantillonnage des objets

La distance de Hausdorff se mesure théoriquement entre deux objets continus. Or la connaissance des objets est échantillonnée. Pour que les distances, entre objets homologues soient représentatives, il faut d'abord que l'échantillonnage respecte l'aspect fréquentiel de la description des contours (au sens de Shannon), et ensuite que les échantillonnages des deux objets comparés soient équivalents. On ramène alors par "rastérisation" les contours vecteurs de la référence au même échantillonnage pixelaire que celui de la segmentation.

Problèmes liés à la différence de spécification des saisies

L'erreur totale, estimée par comparaison à une référence, commise lors de la segmentation, qu'elle soit manuelle ou automatique, est la somme d'une erreur d'échantillonnage géométrique du capteur et des erreurs de saisies.

Les erreurs d'échantillonnage géométrique sont liées au passage de la référence cartographique à sa projection image. Cette projection impose une connaissance de l'échantillonnage

géométrique du capteur qui est trop approximative même après recalage par points d'appui.

Les erreurs de saisies, sont composées d'erreurs intrinsèques des méthodes de segmentation utilisées, d'erreurs de généralisation liées à une différence d'échelle entre les images de travail et les images ayant servi à l'obtention de la référence, et d'erreurs d'interprétations. Les erreurs d'interprétation et généralisation sont liées à la différence entre le terrain vu et le terrain nominal [LEM95]. En d'autres termes, ces erreurs ne sont pas intrinsèques mais liées à la spécification des saisies ou segmentations d'objets.

Pour évaluer la qualité intrinsèque des méthodes de segmentation, il est donc nécessaire de séparer les différentes sources d'erreurs présentes dans les saisies et notamment d'identifier les erreurs liées à la spécification des saisies. L'idée, pour résoudre ce problème influençant la mesure de l'écart entre les primitives image-référence, est d'éliminer ou d'atténuer les influences des détails de saisies qui sont discordants. Cette action appelée "régularisation des contours" va permettre d'éliminer une majeure partie de l'erreur liée à la méthode de comparaison à une référence. [BAD92] propose dans cette optique de modérer automatiquement les discordances par l'atténuation de la distance grâce à une fonction concave. Cette fonction a le même rôle que la boule de tolérance de [HOT95] et [ABB94] qui permet d'éliminer manuellement les détails discordants. Les erreurs restantes après régularisation sont la somme d'erreurs intrinsèques des méthodes de restitution et encore d'erreurs résiduelles de généralisation. Ces erreurs doivent être accompagnées de pourcentage de tracés concordant segmentation-référence et référence-segmentation pour une description significative de la qualité de la segmentation.

Pour estimer les erreurs intrinsèques des méthodes de segmentation, il faut s'affranchir des erreurs liées aux problèmes de généralisation et avoir la référence la plus précise possible. C'est la raison pour laquelle, nous proposons une méthodologie d'évaluation des techniques de segmentations basée sur le concept de la simulation d'images. Ce principe est utilisé à des fins d'auto-validation de la synthèse et de l'analyse d'images dans [GAG90].

La simulation d'images sur des cibles cartographiques simples, non sensibles à la généralisation, permet de s'affranchir de ces problèmes inhérents à la comparaison à une référence et d'avoir une connaissance infinie de la référence.

SIMULATION D'IMAGES REALISTES

AEROSPATIALE a développé un simulateur complet de la chaîne image d'un système optique d'observation de la Terre, permettant de générer des images réalistes non seulement d'un point de vue radiométrique mais prenant également en compte les déformations géométriques induites par la dynamique d'acquisition.

Ce simulateur a été conçu pour l'analyse des performances de qualité image, et est utilisé ici comme outil d'aide à la conception et à la validation des méthodes d'exploitation des images.

La simulation proprement dite, utilise deux types de données sources, les uns décrivant le paysage (relief, albédo), les autres décrivant le système de prise de vue (capteur, trajectoire, attitude, segment sol). [THI92] détaille ce principe de simulation. Dans les deux cas, il est souhaitable que ces données soient aussi paramétriques que possible comme on le montrera plus loin.

Modélisation du paysage

En entrée du simulateur, les données cartographiques sont gérées à deux échelles différentes. Une base de données "basse fréquence" fournit une description du relief sous la forme d'un modèle numérique de terrain, et une information d'albédo sous la forme d'une orthophoto de référence beaucoup plus résolue que l'image à simuler.

Une base de données "haute fréquence" décrit les superstructures cartographiques sous la forme d'un modèle TRAPU fourni par l'IGN et d'un jeu de photographies aériennes sur la même zone. **La forme de ces superstructures est choisie pour échapper le plus possible au problème de la généralisation des formes à travers les échelles.**

La modélisation du paysage consiste à gérer de manière cohérente la relation entre la géométrie et la radiométrie pour chacune des bases de données, et à incruster les bâtiments dans le MNT avec une gestion cohérente des pieds des bâtiments et des ombres portées. La géométrie d'incrustation est choisie de manière paramétrique.

Modélisation du système de prise de vue

Le système de prise de vue est modélisé selon ses aspects géométriques d'une part, radiométriques d'autre part. La modélisation géométrique contient tous les paramètres affectant la localisation (fonction qui à une position dans l'image associe un axe de visée), en particulier: la trajectoire de la plate-forme, l'attitude de la plate-forme, la cartographie du plan focale, la distance focale.

La modélisation radiométrique permet de calculer une radiométrie à partir d'une image de réflectances, d'un modèle d'atmosphère et de la description paramétrique de tous les sous-systèmes affectant la fonction de transfert (FTM) ou la résolution radiométrique.

Enfin, pour chacun des sous-systèmes, on cherche à modéliser une méconnaissance réaliste. C'est particulièrement important pour la validation de techniques de restitution cartographique, qui n'utilisent qu'une connaissance nominale ou estimée du système.

Avantages et contraintes de la simulation d'images

La simulation d'images réaliste présente des avantages multiples pour la validation des techniques de cartographie. D'abord elle permet de multiplier les situations, c'est à dire les paysages d'une part et les systèmes optiques d'autre part. On peut ainsi chercher à éprouver les algorithmes dans des situations un peu extrêmes. Ensuite, cette variété de situations peut être gérée de manière entièrement paramétrique, ce qui permet d'étudier l'impact de chaque paramètre (qu'il soit lié au système ou au terrain, etc) sur les performances obtenues en le faisant varier et en fixant tous les autres.

Malgré les avantages non négligeables cités ci-dessus, les images simulées ont toujours une limite de réalisme qui est nécessairement pénalisante par rapport à de vraies images. Par ailleurs, le besoin de disposer d'un paysage de référence entraîne des contraintes opérationnelles dans la conception algorithmique d'un simulateur. En effet, pour avoir accès au paysage de référence, il est nécessaire de le modéliser indépendamment des conditions d'observations (donc dans l'espace objet et non dans l'espace image)

Validation à partir de simulations

Nous avons vu que la connaissance d'une référence cartographique était rare et qu'une référence issue d'une saisie manuelle à l'échelle de l'image était entachée d'erreurs trop