

Extension de l'étiquetage géographique des pixels d'une image par fouille de données

Adrien Gressin*, Nicole Vincent**, Clément Mallet*, Nicolas Papanoditis*

*IGN/SR, MATIS, Université Paris-Est, Saint-Mandé; (firstname.lastname@ign.fr)

** LIPADE - SIP, Université Paris-Descartes, Paris; (nicole.vincent@mi.parisdescartes.fr)

Résumé. Les techniques de classification modernes permettent d'étiqueter les zones non couvertes des bases de données cartographiques, mais souffrent d'un manque de robustesse important. Dans cet article, nous proposons une méthode robuste d'extension d'étiquetage sur l'emprise d'une image satellite, par analyse hiérarchique des données existantes. Notre approche est fondée sur une sélection d'attributs par thème de la base de données, une sélection des pixels d'apprentissage et des classifications par objet de chaque thème. La décision finale d'étiquetage est prise après fusion des classifications par thème. Notre méthode est appliquée avec succès et comparée à plusieurs méthodes de classification, couplant données d'occupation du sol et imagerie spatiale très haute résolution.

1 Introduction

Dans la plupart des pays développés, des bases de données géographiques (BD) sont initiées, même si certaines régions ne sont couvertes que partiellement. En particulier, des BD d'occupation du sol (OCS) à grande échelle sont en cours de réalisation, par agrégation de données existantes et leur assemblage ne permet pas une description complète du territoire. D'autre part, des images satellite de résolution sub-métrique couvrent avec une grande précision géométrique de larges territoires et peuvent donc aider à compléter ces BD d'OCS. Les méthodes de classification supervisée sont largement utilisées pour résoudre ce genre de problème de télédétection (Mountrakis et al., 2011). Cependant, elles sont souvent peu robustes, présentent un fort taux de confusion, sont limitées à certaines thématiques. Enfin, elles nécessitent une sélection manuelle des classes et des zones d'apprentissage afin de traiter le cas de classes composées de différentes apparences. Nous proposons ainsi ici une méthode d'inspection hiérarchique d'une BD existante, permettant d'apprendre indifféremment la (ou les) apparence(s) de chaque thème qui la constitue. Ces informations sont ensuite fusionnées à différents niveaux afin d'obtenir des résultats plus robustes.

2 Méthodologie

La structure hiérarchique des BD géographiques permet de faire ressortir trois niveaux possibles d'inspection : (1) le niveau objet, (2) le niveau thème, (3) le niveau BD. La BD initiale est projetée sur l'image, formant une carte qui lui est superposable et où les pixels sont