

# Vectorisation de segmentation raster

## Stage d'Initiation à la Recherche

**Mots clef :** Segmentation, vectorisation, optimisation, topologie, géométrie algorithmique.

## 1 Contexte

L'Institut National de l'Information Géographique et Forestière (IGN) a pour mission de produire et de maintenir une représentation fidèle du territoire français. Parmi ces représentations, on peut citer par exemple :

1. L'Occupation du sol à grande échelle, qui donne pour chaque maille d'une grille régulière couvrant la France entière sa typologie (champs, forêt, bâtiment, sol nu, route, ...)
2. Bati3D qui contient une représentation 3D structurée des bâtiments de la plupart des grandes villes françaises.

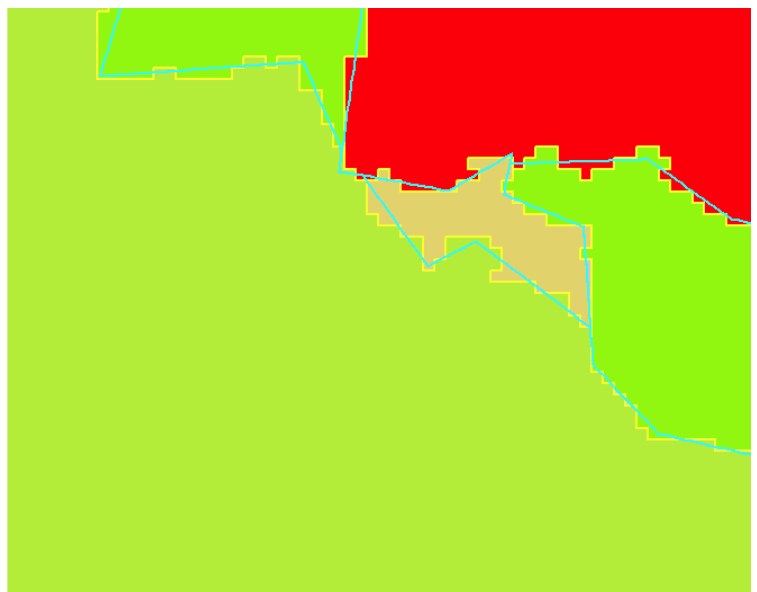
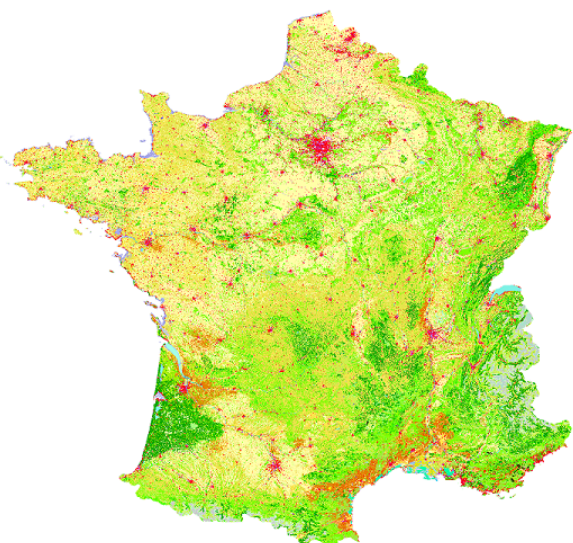
La production de ces deux représentations passe par une segmentation d'une grille régulière, aussi appelée segmentation raster :

1. Pour l'occupation du sol, les mailles correspondant au même objet (même champs, même bâtiment,...) forment des segments.
2. Les bâtiments 3D peuvent être produits en segmentant un Modèle Numérique d'Élévation (MNE) donnant une élévation sur chaque maille d'une grille.

Dans les deux cas, il peut être souhaitable de vectoriser le résultat de la segmentation pour en produire une représentation plus haut niveau et plus utilisable, c'est à dire de produire un polygone fermé représentant chaque segment. Ce type de représentation est très utilisée dans les Systèmes d'Information Géographique (SIG). Pour l'occupation du sol, cette vectorisation n'est pas faite à l'heure actuelle et la représentation raster est utilisée. Pour Bati3D, la vectorisation est faite de façon très heuristique et le résultat est trop généralisé et s'adapte mal aux structures complexes.

## 2 Objectifs du stage

L'objectif de ce projet est de proposer, d'implémenter et de tester une méthodologie générale de vectorisation d'une segmentation raster. Celle-ci devra être faite en tenant compte des trois objectifs suivants :



L'occupation du sol sur la France entière (gauche) et un gros plan sur une petite zone (droite). La segmentation raster induite est représentée par ses contours (jaune) et une vectorisation de ces contours est donnée en bleu clair. Source : Union européenne - SOeS, CORINE Land Cover

1. Fidélité : la vectorisation doit être la plus proche possible de la segmentation raster
2. Généralisation : la vectorisation doit être la plus simple possible (avoir le moins de sommets de polygones possibles)
3. Topologie : la segmentation doit être topologiquement consistante, c'est à dire que deux polygones adjacents doivent partager exactement une partie de leurs faces, et que ces faces ne doivent pas s'auto-intersecter.

Les deux premiers objectifs sont contradictoires et il faudra donc proposer le meilleur compromis possible par l'intermédiaire d'une optimisation. Le troisième objectif apporte une contrainte qui peut être difficile à prendre en compte dans cette optimisation. Les heuristiques classiques de polygonalisation [Douglas73] ne donnent pas cette garantie, et elle est difficile à appliquer à posteriori. Ainsi, on propose de se tourner vers une approche itérative de rectification des contours où la topologie est vérifiée à chaque étape par un algorithme cinétique permettant de détecter et de traiter dans l'ordre tous les événements topologiques (collisions de contours) qui peuvent survenir. Il faudra donc proposer un traitement ad-hoc pour ces collisions permettant de préserver à la fois une topologie correcte et simple. Enfin, le stage s'intéressera à la problématique du passage à l'échelle car la méthode développée a vocation à être utilisée sur des territoires vastes à l'échelle d'une région, d'un pays, voire du monde entier.

### 3 Bibliographie

[Douglas73] David Douglas, Thomas Peucker, "Algorithms for the reduction of the number of points required to represent a digitized line or its caricature", *The Canadian Cartographer* 10(2), 112-122 (1973)

[Lecot06] Gregory Lecot and Bruno Lévy. "ARDECO : Automatic Region DETection and Conversion". Eurographics Symposium on Rendering, 2006

[Mena06] J. B. Mena. "Automatic vectorization of segmented road networks by geometrical and topological analysis of high resolution binary images". *Journal of Knowledge-Based Systems*, Volume 19 Issue 8, December, 2006. Pages 704-718

[Vallet09] B. Vallet, P. Pierrot-Deseilligny, D. Boldo. Building Footprint Database Improvement for 3D reconstruction : a Direction Aware Split and Merge Approach. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*. Vol. 38 (Part 3/W4), pp.139-144, Paris, France, September 2009

### 4 Environnement technique

L'environnement de développement est laissé libre, mais le C++ est imposé comme langage de programmation.

### 5 Profil attendu

- Elève ingénieur ou master 2 recherche
- Spécialisation en traitement du signal et de l'image, vision par ordinateur, machine learning ou optimisation.
- Bonne maîtrise de MatLab ou du langage C++
- Autonomie, rigueur, pragmatisme
- Goût prononcé pour la recherche scientifique

### 6 Accueil, encadrement et contact

Le stage sera réalisé dans l'équipe MATIS<sup>1</sup> du laboratoire LaSTIG de l'IGN (Institut National de l'Information Géographique et Forestière), situé à Saint-Mandé (limitrophe Paris, ligne 1 du métro, station Saint-Mandé). L'équipe MATIS regroupe une trentaine de chercheurs spécialistes des domaines de la vision par ordinateur, photogrammétrie et télédétection, pour l'image terrestre, aérienne et satellite multi-sources. Les encadrants et personnes à contacter sont :

	Bruno VALLET	Mathieu BREDIF
<u>Téléphone</u> :	01 43 98 80 81	01 43 98 83 19
<u>Courriel</u> :	bruno.vallet@ign.fr	mathieu.bredif@ign.fr

---

1. Site web du MATIS : <http://recherche.ign.fr/labos/matis/accueilMATIS.php>